This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.



us C3>

-CHLAMYDIA PNEUMONIAE-GENOME SEQUENCE -

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application is related to 60/128,606, filed April 8, 1999 and 60/108,279, filed November 12, 1998, which are incorporated herein by reference.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

10

FIELD OF THE INVENTION

This invention relates to nucleic acids and polypeptides from *Chlamydia* pneumoniae and to their use in the diagnosis, prevention and treatment of diseases associated with *C. pneumoniae*.

15

20

25

BACKGROUND OF THE INVENTION

Chlamydiaceae is a family of obligate intracellular parasite with a tropism for epithelial cells lining the mucus membranes. The bacteria have two morphologically distinct forms, "elementary body" and "reticulate body". The elementary body is the infectious form, and has a rigid cell wall, primarily of cross-linked outer membrane proteins. The reticulate body is the intracellular, metabolically active form. A unique developmental cycle between these two forms characterizes Chlamydia growth.

C. pneumoniae is a human respiratory pathogen that causes acute respiratory disease, and approximately 10% of community-acquired pneumonia. Antibody prevalence studies have shown that virtually everyone is infected with C. pneumoniae at some time, and that reinfection is common. In addition to respiratory disease, studies have shown an association of this organism with coronary artery disease. It has been demonstrated in atherosclerotic lesions of the aorta and coronary arteries by immunocytochemistry and by polymerase chain reaction (Kuo et al. (1993) J Infect Dis 167(4):841-849).

30

Recent reports have further demonstrated the presence of *C. pneumoniae* in the walls of abdominal aortic aneurysms (Juvonen *et al.* (1997) <u>J Vasc Surg</u> **25**(3):499-505). Abdominal aortic aneurysms are frequently associated with atherosclerosis, and inflammation may be an important factor in aneurysmal dilatation.

10

15

20

25

30



C. pneumoniae may play a role in maintaining an inflammation and triggering the development of aortic aneurysms.

Muhlestein et al. (1996) JACC 27:1555-61, reported a differential incidence of *Chlamydia* species within the coronary artery wall of patients with atherosclerosis versus those with other forms of cardiovascular disease. The extremely high rate of possible infection in patients with symptomatic atherosclerotic disease compared to the very low rate in patients with normal coronary arteries or coronary artery disease from chronic transplant rejection provides evidence for a direct link between the atherosclerotic process and *Chlamydia* infection. Because a history of chlamydial infection is so prevalent in the population, the issue of causality remains. On a physiologic and pathologic level, abnormal interactions among endothelial cells, platelets, macrophages and lymphocytes may lead to a cascade of events resulting in acute endothelial damage, thrombosis and repair, chronically leading to the development of atheroma in blood vessels.

C. pneumoniae is related to other Chlamydia species, but the level of sequence similarity is relatively low. Very little is known about the biology of this organism, although it appears to be an important human pathogen. Allelic diversity and structural relationships between specific genes of Chlamydial species is described in Kaltenboeck et al. (1993) J Bacteriol 175(2):487-502; Gaydos et al. (1992) Infect Immun 60(12):5319-5323; Everett et al. (1997) Int J Syst Bacteriol 47(2):461-473; and Pudjiatmoko et al. (1997) Int J Syst Bacteriol 47(2):425-431.

A number of studies have been published describing methods for detection of *C. pneumoniae*, and for distinguishing between Chlamydial species. Such methods include PCR detection (Rasmussen *et al.* (1992) Mol Cell Probes 6(5):389-394; Holland *et al.* (1990) J Infect Dis 162(4):984-987); a simplified polymerase chain reaction-enzyme immunoassay (Wilson et al. (1996) J Appl Bacteriol 80(4):431-438); sequence determination and restriction endonuclease cleavage (Herrmann *et al.* (1996) J Clin Microbiol 34(8):1897-1902).

Antigenic and molecular analyses of different *C. pneumoniae* strains is described in Jantos *et al.* (1997) <u>J Clin Microbiol</u> 35(3):620-623. Some genes of *C. pneumoniae* have been isolated and sequenced. These include the Gro E operon (Kikuta et al. (1991) <u>Infect Immun</u> 59(12):4665-4669); the major outer membrane protein Perez *et*

al. (1991) Infect Immun 59(6):2195-2199; the DnaK protein homolog (Kornak et al. (1991) Infect Immun 59(2):721-725); as well as a number of ribosomal and other genes.

5

SUMMARY OF THE INVENTION

This invention provides the genomic sequence of *Chlamydia pneumoniae*. The sequence information is useful for a variety of diagnostic and analytical methods. The genomic sequence may be embodied in a variety of media, including computer readable forms, or as a nucleic acid comprising a selected fragment of the sequence. Such fragments generally consist of an open reading frame, transcriptional or translational control elements, or fragments derived therefrom. Proteins encoded by the open reading frames are useful for diagnostic purposes, as well as for their enzymatic or structural activity.

15

20

25

10

DEFINITIONS

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group., e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

30

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

10

15

20

25

30

"Amplification" primers are oligonucleotides comprising either natural or analogue nucleotides that can serve as the basis for the amplification of a select nucleic acid sequence. They include, e.g., polymerase chain reaction primers and ligase chain reaction oligonucleotides.

"Antibody" refers to an immunoglobulin molecule able to bind to a specific epitope on an antigen. Antibodies can be a polyclonal mixture or monoclonal. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Antibodies may exist in a variety of forms including, for example, Fv, F_{ab}, and F(ab)₂, as well as in single chains. Single-chain antibodies, in which genes for a heavy chain and a light chain are combined into a single coding sequence, may also be used.

An "antigen" is a molecule that is recognized and bound by an antibody, e.g., peptides, carbohydrates, organic molecules, or more complex molecules such as glycolipids and glycoproteins. The part of the antigen that is the target of antibody binding is an antigenic determinant and a small functional group that corresponds to a single antigenic determinant is called a hapten.

"Biological sample" refers to any sample obtained from a living or dead organism. Examples of biological samples include biological fluids and tissue specimens. Such biological samples can be prepared for analysis of the presence of *C. pneumoniae* nucleic acids, proteins, or antibodies specifically reactive with the proteins.

The term "C. pneumoniae gene" shall be intended to mean the open reading frame encoding specific C. pneumoniae polypeptides, as well as adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 2 kb beyond the coding region, but possibly further in either direction. The gene may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues

5

10

(Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silen: variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following groups each contain amino acids that are conservative substitutions for one another:

25 1) Alanine (A), Glycine (G); Serine (S), Threonine (T); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) 5) Cysteine (C), Methionine (M); 30 6) Arginine (R), Lysine (K), Histidine (H); 7) Isoleucine (I), Leucine (L), Valine (V); and Phenylalanine (F), Tyrosine (Y), Tryptophan (W). 8) see, e.g., Creighton, Proteins (1984)).

15

10

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence, which has a designated percent sequence or subsequence complementarity when the test sequence has a designated or substantial identity to a reference sequence. For example, a designated amino acid percent identity of 95% refers to sequences or subsequences that have at least about 95% amino acid identity when aligned for maximum correspondence over a comparison window as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences would then be said to have substantial identity, or to be substantially identical to each other. Preferably, sequences have at least about 70% identity, more preferably 80% identity, more preferably 90-95% identity and above. Preferably, the percent identity exists over a region of the sequence that is at least about 25 amino acids in length, more preferably over a region that is 50-100 amino acids in length.

When percentage of sequence identity is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acids residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for 25 making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of 30 conservative substitutions is calculated according to, e.g., the algorithm of Meyers & Miller, Computer Applic. Biol. Sci. 4:11-17 (1988) e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California, USA)..

15

20

25

30

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated or default program parameters.

A comparison window includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 25 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel et al., supra).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The

25

30

5

final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, e.g, version 7.0 (Devereaux et al., Nuc. Acids Res. 12:387-395 (1984).

Another example of algorithm that is suitable for determining percent sequence identity (i.e., substantial similarity or identity) is the BLAST algorithm, which 10 is described in Altschul et al., J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues, always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as default parameters a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)).

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below.

Another indication that polynucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically stringent conditions for a Southern blot protocol involve hybridizing in a buffer comprising 5x SSC, 1% SDS at 65°C or hybridizing in a buffer containing 5x SSC and 1% SDS at 42°C and washing at 65°C with a 0.2x SSC, 0.1% SDS wash.

A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available.

15

20

25

30

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

A labeled nucleic acid probe or oligonucleotide is one that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

5

10

15

20

25

30

"Pharmaceutically acceptable" means a material that is not biologically or otherwise undesirable, i.e., the material can be administered to an individual along with a *Chlamydia* antigen without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an analog or mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers.

The phrase "specifically or selectively hybridizing to," refers to hybridization between a probe and a target sequence in which the probe binds substantially only to the target sequence, forming a hybridization complex, when the target is in a heterogeneous mixture of polynucleotides and other compounds. Such hybridization is determinative of the presence of the target sequence. Although the probe may bind other unrelated sequences, at least 90%, preferably 95% or more of the hybridization complexes formed are with the target sequence.

The term "recombinant" when used with reference to a cell, or nucleic acid, or vector, indicates that the cell, or nucleic acid, or vector, has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The phrase "specifically immunoreactive with", when referring to a protein or peptide, refers to a binding reaction between the protein and an antibody which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other compounds. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein and are described in detail below.

The phrase "substantially pure" or "isolated" when referring to a Chlamydia peptide or protein, means a chemical composition which is free of other subcellular components of the Chlamydia organism. Typically, a monomeric protein is substantially pure when at least about 85% or more of a sample exhibits a single polypeptide backbone. Minor variants or chemical modifications may typically share the same polypeptide sequence. Depending on the purification procedure, purities of 85%, and preferably over 95% pure are possible. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band on a polyacrylamide gel upon silver staining. For certain purposes high resolution will be needed and HPLC or a similar means for purification utilized.

10

DETAILED DESCRIPTION

The present invention provides the nucleotide sequence of the C. pneumoniae genome SEQ ID NO: 1 or a representative fragment thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan. As used herein, a "representative fragment" of the nucleotide sequence depicted in SEQ ID NO: 1 refers to any portion which is not presently represented within a publicly available database. Preferred representative fragments of the present invention are open reading frames, expression modulating fragments, uptake modulating fragments, and fragments which can be used to diagnose the presence of C. pneumoniae in sample. Using the information provided in the present application, together with routine cloning and sequencing methods, one of ordinary skill in the art will be able to clone and sequence all "representative fragments" of interest including open reading frames (ORFs) encoding a large variety of C. pneumoniae proteins. A non-limiting identification of such preferred representative fragments is provided in Tables 2 and 2.

Ons 25 A'

Diagnostic use of C. pneumoniae nucleic acids

Hybridization-based assays

Using the nucleic acids disclosed here, one of skill can design nucleic acid hybridization-based assays for the detection of *C. pneumoniae*. Any of a number of well known techniques for the specific detection of target nucleic acids can be used.

Exemplary hybridization-based assays include, but are not limited to, traditional "direct

20

25

30

5

10

probe" methods such as Southern Blots, dot blots, in situ hybridization (e.g., FISH), PCR, and the like. The methods can be used in a wide variety of formats including, but not limited to substrate- (e.g. membrane or glass) bound methods or array-based approaches as described below. As noted above, this invention also embraces methods for detecting the presence of Chlamydia DNA or RNA in biological samples. These sequences can be used to detect Chlamydia in biological samples from patients suspected of being infected. A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art (see Sambrook et al., supra).

In situ hybridization assays are well known (e.g., Angerer (1987) Meth. Enzymol 152: 649). Generally, in situ hybridization comprises the following major steps: (1) fixation of tissue or biological structure to analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments. The reagent used in each of these steps and the conditions for use vary depending on the particular application.

In a typical *in situ* hybridization assay, cells are fixed to a solid support, typically a glass slide. If a nucleic acid is to be probed, the cells are typically denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of labeled probes specific to the nucleic acid sequence encoding the protein. The targets (e.g., cells) are then typically washed at a predetermined stringency or at an increasing stringency until an appropriate signal to noise ratio is obtained.

The nucleic acids of this invention are particularly well suited to array-based hybridization formats. Arrays are a multiplicity of different "probe" or "target" nucleic acids (or other compounds) attached to one or more surfaces (e.g., solid, membrane, or gel). In a preferred embodiment, the multiplicity of nucleic acids (or other moieties) is attached to a single contiguous surface or to a multiplicity of surfaces juxtaposed to each other.

In an array format a large number of different hybridization reactions can be run essentially "in parallel." This provides rapid, essentially simultaneous, evaluation

10

15

20

of a number of hybridizations in a single "experiment". Methods of performing hybridization reactions in array based formats are well known to those of skill in the art (see, e.g., Pastinen (1997) Genome Res. 7: 606-614; Jackson (1996) Nature Biotechnology 14:1685; Chee (1995) Science 274: 610; WO 96/17958.

Arrays, particularly nucleic acid arrays can be produced according to a wide variety of methods well known to those of skill in the art. For example, in a simple embodiment, "low density" arrays can simply be produced by spotting (e.g. by hand using a pipette) different nucleic acids at different locations on a solid support (e.g. a glass surface, a membrane, etc.).

This simple spotting, approach has been automated to produce high density spotted arrays (*see*, *e.g.*, U.S. Patent No: 5,807,522). This patent describes the use of an automated systems that taps a microcapillary against a surface to deposit a small volume of a biological sample. The process is repeated to generate high density arrays. Arrays can also be produced using oligonucleotide synthesis technology. Thus, for example, U.S. Patent No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092 teach the use of light-directed combinatorial synthesis of high density oligonucleotide arrays.

Many methods for immobilizing nucleic acids on a variety of solid surfaces are known in the art. A wide variety of organic and inorganic polymers, as well as other materials, both natural and synthetic, can be employed as the material for the solid surface. Illustrative solid surfaces include, e.g., nitrocellulose, nylon, glass, quartz, diazotized membranes (paper or nylon), silicones, polyformaldehyde, cellulose, and cellulose acetate. In addition, plastics such as polyethylene, polypropylene, polystyrene, and the like can be used. Other materials which may be employed include paper, ceramics, metals, metalloids, semiconductive materials, cermets or the like. In addition, substances that form gels can be used. Such materials include, e.g., proteins (e.g., gelatins), lipopolysaccharides, silicates, agarose and polyacrylamides. Where the solid surface is porous, various pore sizes may be employed depending upon the nature of the system.

In preparing the surface, a plurality of different materials may be employed, particularly as laminates, to obtain various properties. For example, proteins (e.g., bovine serum albumin) or mixtures of macromolecules (e.g., Denhardt's solution) can be employed to avoid non-specific binding, simplify covalent conjugation, enhance

14

30

25

5

10

15

20

25

signal detection or the like. If covalent bonding between a compound and the surface is desired, the surface will usually be polyfunctional or be capable of being polyfunctionalized. Functional groups which may be present on the surface and used for linking can include carboxylic acids, aldehydes, amino groups, cyano groups, ethylenic groups, hydroxyl groups, mercapto groups and the like. The manner of linking a wide variety of compounds to various surfaces is well known and is amply illustrated in the literature.

For example, methods for immobilizing nucleic acids by introduction of various functional groups to the molecules is known (see, e.g., Bischoff (1987) Anal. Biochem., 164: 336-344; Kremsky (1987) Nucl. Acids Res. 15: 2891-2910). Modified nucleotides can be placed on the target using PCR primers containing the modified nucleotide, or by enzymatic end labeling with modified nucleotides. Use of glass or membrane supports (e.g., nitrocellulose, nylon, polypropylene) for the nucleic acid arrays of the invention is advantageous because of well developed technology employing manual and robotic methods of arraying targets at relatively high element densities. Such membranes are generally available and protocols and equipment for hybridization to membranes is well known.

Target elements of various sizes, ranging from 1 mm diameter down to 1 μm can be used. Smaller target elements containing low amounts of concentrated, fixed probe DNA are used for high complexity comparative hybridizations since the total amount of sample available for binding to each target element will be limited. Thus it is advantageous to have small array target elements that contain a small amount of concentrated probe DNA so that the signal that is obtained is highly localized and bright. Such small array target elements are typically used in arrays with densities greater than 10⁴/cm². Relatively simple approaches capable of quantitative fluorescent imaging of 1 cm² areas have been described that permit acquisition of data from a large number of target elements in a single image (see, e.g., Wittrup (1994) Cytometry 16:206-213).

If fluorescently labeled nucleic acid samples are used, arrays on solid surface substrates with much lower fluorescence than membranes, such as glass, quartz, or small beads, can achieve much better sensitivity. Substrates such as glass or fused silica are advantageous in that they provide a very low fluorescence substrate, and a highly efficient hybridization environment. Covalent attachment of the target nucleic acids to glass or synthetic fused silica can be accomplished according to a number of

10

15

20

25

30

known techniques (described above). Nucleic acids can be conveniently coupled to glass using commercially available reagents. For instance, materials for preparation of silanized glass with a number of functional groups are commercially available or can be prepared using standard techniques (see, e.g., Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Wash., D.C.). Quartz cover slips, which have at least 10-fold lower autofluorescence than glass, can also be silanized.

Alternatively, probes can also be immobilized on commercially available coated beads or other surfaces. For instance, biotin end-labeled nucleic acids can be bound to commercially available avidin-coated beads. Streptavidin or anti-digoxigenin antibody can also be attached to silanized glass slides by protein-mediated coupling using e.g., protein A following standard protocols (see, e.g., Smith (1992) Science 258: 1122-1126). Biotin or digoxigenin end-labeled nucleic acids can be prepared according to standard techniques. Hybridization to nucleic acids attached to beads is accomplished by suspending them in the hybridization mix, and then depositing them on the glass substrate for analysis after washing. Alternatively, paramagnetic particles, such as ferric oxide particles, with or without avidin coating, can be used.

A variety of other nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in Hames and Higgins (1985) *Nucleic Acid Hybridization, A Practical Approach*, IRL Press; Gall and Pardue (1969) *Proc. Natl. Acad. Sci. USA* 63: 378-383; and John *et al.* (1969) *Nature* 223: 582-587.

Sandwich assays are commercially useful hybridization assays for detecting or isolating nucleic acid sequences. Such assays utilize a "capture" nucleic acid covalently immobilized to a solid support and a labeled "signal" nucleic acid in solution. The sample will provide the target nucleic acid. The "capture" nucleic acid and "signal" nucleic acid probe hybridize with the target nucleic acid to form a "sandwich" hybridization complex. To be most effective, the signal nucleic acid should not hybridize with the capture nucleic acid.

Detection of a hybridization complex may require the binding of a signal generating complex to a duplex of target and probe polynucleotides or nucleic acids.

Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal.

10

15

20

25

30

The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system that multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction (PCR) system and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBAO, Cangene, Mississauga, Ontario) and Q Beta Replicase systems.

Nucleic acid hybridization simply involves providing a denatured probe and target nucleic acid under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids, or in the addition of chemical agents, or the raising of the pH. Under low stringency conditions (e.g., low temperature and/or high salt and/or high target concentration) hybrid duplexes (e.g., DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (e.g., higher temperature or lower salt) successful hybridization requires fewer mismatches.

One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency. In a preferred embodiment, hybridization is performed at low stringency to ensure hybridization and then subsequent washes are performed at higher stringency to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency (e.g., down to as low as 0.25 X SSPE-T at 37°C to 70°C) until a desired level of hybridization specificity is obtained. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that can be present.

In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in a preferred embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, in a preferred embodiment, the hybridized array may be washed at successively higher

10

15

20

25

30

stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular probes of interest.

Methods of optimizing hybridization conditions are well known to those of skill in the art (see, e.g., Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes, Elsevier, N.Y.).

Labeling and detection of nucleic acids.

In a preferred embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample or probe nucleic acids. The labels may be incorporated by any of a number of means well known to those of skill in the art. Means of attaching labels to nucleic acids include, for example nick translation or endlabeling (e.g. with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (e.g., a fluorophore). A wide variety of linkers for the attachment of labels to nucleic acids are also known. In addition, intercalating dyes and fluorescent nucleotides can also be used.

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (*e.g.*, DynabeadsTM), fluorescent dyes (*e.g.*, fluorescein, texas red, rhodamine, green fluorescent protein, and the like, *see*, *e.g.*, Molecular Probes, Eugene, Oregon, USA), radiolabels (*e.g.*, ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold (*e.g.*, gold particles in the 40 -80 nm diameter size range scatter green light with high efficiency) or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

A fluorescent label is preferred because it provides a very strong signal with low background. It is also optically detectable at high resolution and sensitivity through a quick scanning procedure. The nucleic acid samples can all be labeled with a single label, e.g., a single fluorescent label. Alternatively, in another embodiment, different nucleic acid samples can be simultaneously hybridized where each nucleic acid

5

10

15

20

25

30

sample has a different label. For instance, one target could have a green fluorescent label and a second target could have a red fluorescent label. The scanning step will distinguish cites of binding of the red label from those binding the green fluorescent label. Each nucleic acid sample (target nucleic acid) can be analyzed independently from one another.

Suitable chromogens which can be employed include those molecules and compounds which absorb light in a distinctive range of wavelengths so that a color can be observed or, alternatively, which emit light when irradiated with radiation of a particular wave length or wave length range, e.g., fluorescers.

Desirably, fluorescers should absorb light above about 300 nm, preferably about 350 nm, and more preferably above about 400 nm, usually emitting at wavelengths greater than about 10 nm higher than the wavelength of the light absorbed. It should be noted that the absorption and emission characteristics of the bound dye can differ from the unbound dye. Therefore, when referring to the various wavelength ranges and characteristics of the dyes, it is intended to indicate the dyes as employed and not the dye which is unconjugated and characterized in an arbitrary solvent.

Fluorescers are generally preferred because by irradiating a fluorescer with light, one can obtain a plurality of emissions. Thus, a single label can provide for a plurality of measurable events.

Detectable signal can also be provided by chemiluminescent and bioluminescent sources. Chemiluminescent sources include a compound which becomes electronically excited by a chemical reaction and can then emit light which serves as the detectable signal or donates energy to a fluorescent acceptor. Alternatively, luciferins can be used in conjunction with luciferase or lucigenins to provide bioluminescence. Spin labels are provided by reporter molecules with an unpaired electron spin which can be detected by electron spin resonance (ESR) spectroscopy. Exemplary spin labels include organic free radicals, transitional metal complexes, particularly vanadium, copper, iron, and manganese, and the like. Exemplary spin labels include nitroxide free radicals.

The label may be added to the target (sample) nucleic acid(s) prior to, or after the hybridization. So called "direct labels" are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, so called "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the

10

15

25

target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected. For a detailed review of methods of labeling nucleic acids and detecting labeled hybridized nucleic acids see Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes, P. Tijssen, ed. Elsevier, N.Y., (1993)).

Fluorescent labels are easily added during an in vitro transcription reaction. Thus, for example, fluorescein labeled UTP and CTP can be incorporated into the RNA produced in an in vitro transcription.

The labels can be attached directly or through a linker moiety. In general, the site of label or linker-label attachment is not limited to any specific position. For example, a label may be attached to a nucleoside, nucleotide, or analogue thereof at any position that does not interfere with detection or hybridization as desired. For example, certain Label-ON Reagents from Clontech (Palo Alto, CA) provide for labeling interspersed throughout the phosphate backbone of an oligonucleotide and for terminal labeling at the 3' and 5' ends. As shown for example herein, labels can be attached at positions on the ribose ring or the ribose can be modified and even eliminated as desired. The base moieties of useful labeling reagents can include those that are naturally occurring or modified in a manner that does not interfere with the purpose to which they are put. Modified bases include but are not limited to 7-deaza A and G, 7-deaza-8-aza A and G, and other heterocyclic moieties.

It will be recognized that fluorescent labels are not to be limited to single species organic molecules, but include inorganic molecules, multi-molecular mixtures of organic and/or inorganic molecules, crystals, heteropolymers, and the like. Thus, for example, CdSe-CdS core-shell nanocrystals enclosed in a silica shell can be easily derivatized for coupling to a biological molecule (Bruchez et al. (1998) Science, 281: 2013-2016). Similarly, highly fluorescent quantum dots (zinc sulfide-capped cadmium selenide) have been covalently coupled to biomolecules for use in ultrasensitive biological detection (Warren and Nie (1998) Science, 281: 2016-2018).

30

Amplification-based assays.

In another embodiment, amplification-based assays can be used to detect nucleic acids. In such amplification-based assays, the nucleic acid sequences act as a

10

15

20

template in an amplification reaction (e.g. Polymerase Chain Reaction (PCR). Detailed protocols for quantitative PCR are provided in Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, Inc. N.Y.).

Other suitable amplification methods include, but are not limited to ligase chain reaction (LCR) (see Wu and Wallace (1989) Genomics 4: 560, Landegren et al. (1988) Science 241: 1077, and Barringer et al. (1990) Gene 89: 117, transcription amplification (Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173), and self-sustained sequence replication (Guatelli et al. (1990) Proc. Nat. Acad. Sci. USA 87: 1874).

Detection of C. pneumoniae gene expression

The nucleic acids of the invention can also be used to *C. pneumoniae* detect gene transcripts. Methods of detecting and/or quantifying gene transcripts using nucleic acid hybridization techniques are known to those of skill in the art (see Sambrook *et al. supra*). For example, a Northern transfer may be used for the detection of the desired mRNA directly. In brief, the mRNA is isolated from a given cell sample using, for example, an acid guanidinium-phenol-chloroform extraction method. The mRNA is then electrophoresed to separate the mRNA species and the mRNA is transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled probes are used to identify and/or quantify the target mRNA.

In another preferred embodiment, the gene transcript can be measured using amplification (e.g. PCR) based methods as described above for directly assessing copy number of the target sequences.

Expression of C. pneumoniae proteins

The nucleic acids disclosed here can be used for recombinant expression

of the proteins. In these methods, the nucleic acids encoding the proteins of interest are introduced into suitable host cells, followed by induction of the cells to produce large amounts of the protein. The invention relies on routine techniques in the field of recombinant genetics, well known to those of ordinary skill in the art. A basic text disclosing the general methods of use in this invention is Sambrook et al., Molecular

Cloning, A Laboratory Manual (2nd ed. 1989).

Standard transfection methods are used to produce prokaryotic, mammalian, yeast or insect cell lines which express large quantities of the desired

15

20

25

30

polypeptide, which is then purified using standard techniques (see, e.g., Colley et al., J. Biol. Chem. 264:17619-17622, 1989; Guide to Protein Purification, supra).

The nucleotide sequences used to transfect the host cells can be modified to yield *Chlamydia* polypeptides with a variety of desired properties. For example, the polypeptides can vary from the naturally-occurring sequence at the primary structure level by amino acid, insertions, substitutions, deletions, and the like. These modifications can be used in a number of combinations to produce the final modified protein chain.

The amino acid sequence variants can be prepared with various objectives in mind, including facilitating purification and preparation of the recombinant polypeptide. The modified polypeptides are also useful for modifying plasma half life, improving therapeutic efficacy, and lessening the severity or occurrence of side effects during therapeutic use. The amino acid sequence variants are usually predetermined variants not found in nature but exhibit the same immunogenic activity as naturally occurring protein. In general, modifications of the sequences encoding the polypeptides may be readily accomplished by a variety of well-known techniques, such as site-directed mutagenesis (see Gillman & Smith, Gene 8:81-97 (1979); Roberts et al., Nature 328:731-734 (1987)). One of ordinary skill will appreciate that the effect of many mutations is difficult to predict. Thus, most modifications are evaluated by routine screening in a suitable assay for the desired characteristic. For instance, the effect of various modifications on the ability of the polypeptide to elicit a protective immune response can be easily determined using in vitro assays. For instance, the polypeptides can be tested for their ability to induce lymphoproliferation, T cell cytotoxicity, or cytokine production using standard techniques.

The particular procedure used to introduce the genetic material into the host cell for expression of the polypeptide is not particularly critical. Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasmid vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see Sambrook et al., supra). It is only necessary that the particular procedure utilized be capable of successfully introducing at least one gene into the host cell which is capable of expressing the gene.

20

25

30

5

10

Any of a number of well known cells and cell lines can be used to express the polypeptides of the invention. For instance, prokaryotic cells such as *E. coli* can be used. Eukaryotic cells include, yeast, Chinese hamster ovary (CHO) cells, COS cells, and insect cells.

The particular vector used to transport the genetic information into the cell is also not particularly critical. Any of the conventional vectors used for expression of recombinant proteins in prokaryotic and eukaryotic cells may be used. Expression vectors for mammalian cells typically contain regulatory elements from eukaryotic viruses.

The expression vector typically contains a transcription unit or expression cassette that contains all the elements required for the expression of the polypeptide DNA in the host cells. A typical expression cassette contains a promoter operably linked to the DNA sequence encoding a polypeptide and signals required for efficient polyadenylation of the transcript. The term "operably linked" as used herein refers to linkage of a promoter upstream from a DNA sequence such that the promoter mediates transcription of the DNA sequence. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

Following the growth of the recombinant cells and expression of the polypeptide, the culture medium is harvested for purification of the secreted protein. The media are typically clarified by centrifugation or filtration to remove cells and cell debris and the proteins are concentrated by adsorption to any suitable resin or by use of ammonium sulfate fractionation, polyethylene glycol precipitation, or by ultrafiltration. Other routine means known in the art may be equally suitable. Further purification of the polypeptide can be accomplished by standard techniques, for example, affinity chromatography, ion exchange chromatography, sizing chromatography, His₆ tagging and Ni-agarose chromatography (as described in Dobeli *et al.*, *Mol. and Biochem. Parasit.* 41:259-268 (1990)), or other protein purification techniques to obtain homogeneity. The purified proteins are then used to produce pharmaceutical compositions, as described below.

An alternative method of preparing recombinant polypeptides useful as vaccines involves the use of recombinant viruses (e.g., vaccinia). Vaccinia virus is grown

5

10

15

20

25

30

in suitable cultured mammalian cells such as the HeLa S3 spinner cells, as described by Mackett et al., in DNA cloning Vol. II: A practical approach, pp. 191-211 (Glover, ed.).

Antibody Production

The proteins of the present invention can be used to produce antibodies specifically reactive with *C pneumoniae* antigens. If isolated proteins are used, they may be recombinantly produced or isolated from *Chlamydia* cultures. Synthetic peptides made using the protein sequences may also be used.

Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera is prepared. Further fractionation of the antisera to enrich for antibodies reactive to *Chlamydia* proteins can be done if desired (see Harlow & Lane, *Antibodies: A Laboratory Manual* (1988)).

Polyclonal antisera are used to identify and characterize *Chlamydia* in the tissues of patients using, for instance, *in situ* techniques and immunoperoxidase test procedures described in Anderson *et al. JAVMA* 198:241 (1991) and Barr *et al. Vet. Pathol.* 28:110-116 (1991).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see Kohler & Milstein, Eur. J. Immunol. 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Monoclonal antibodies produced in such a manner are used, for instance, in ELISA diagnostic tests, immunoperoxidase tests, immunohistochemical tests, for the *in vitro* evaluation of spirochete invasion, to select candidate antigens for vaccine development, protein isolation, and for screening genomic and cDNA libraries to select appropriate gene sequences.

5

10

15

20

25

30

Immunodiagonostic detection of C. pneumoniae infections

The present invention also provides methods for detecting the presence or absence of *C. pneumoniae*, or antibodies reactive with it, in a biological sample. For instance, antibodies specifically reactive with *Chlamydia* can be detected *using* either *Chlamydia* proteins or the isolates described here. The proteins and isolates can also be used to raise specific antibodies (either monoclonal or polyclonal) to detect the antigen in a sample. In addition, the nucleic acids disclosed and claimed here can be used to detect *Chlamydia*-specific sequences using standard hybridization techniques.

For a review of immunological and immunoassay procedures in general, see *Basic and Clinical Immunology* (Stites & Terr ed., 7th ed. 1991)). The immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio, ed., 1980); Tijssen, *Laboratory Techniques in Biochem.stry and Molecular Biology* (1985)). For instance, the proteins and antibodies disclose I here are conveniently used in ELISA, immunoblot analysis and agglutination assays.

In brief, immunoassays to measure anti-Chlamydia antibodies or antigens can be either competitive or noncompetitive binding assays. In competitive binding assays, the sample analyte (e.g., anti-Chlamydia antibodies) competes with a labeled analyte (e.g., anti-Chlamydia monoclonal antibody) for specific binding sites on a capture agent (e.g., isolated Chlamydia protein) bound to a solid surface. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

Noncompetitive assays are typically sandwich assays, in which the sample analyte is bound between two analyte-specific binding reagents. One of the binding agents is used as a capture agent and is bound to a solid surface. The second binding agent is labelled and is used to measure or detect the resultant complex by visual or instrument means.

A number of combinations of capture agent and labelled binding agent can be used. For instance, an isolated *Chlamydia* protein or culture can be used as the capture agent and labelled anti-human antibodies specific for the constant region of human antibodies can be used as the labelled binding agent. Goat, sheep and other non-human antibodies specific for human immunoglobulin constant regions (e.g., γ or μ) are

5

10

15

20

25

well known in the art. Alternatively, the anti-human antibodies can be the capture agent and the antigen can be labelled.

Various components of the assay, including the antigen, anti-Chlamydia antibody, or anti-human antibody, may be bound to a solid surface. Many methods for immobilizing biomolecules to a variety of solid surfaces are known in the art. For instance, the solid surface may be a membrane (e.g., nitrocellulose), a microtiter dish (e.g., PVC or polystyrene) or a bead. The desired component may be covalently bound or noncovalently attached through nonspecific bonding.

Alternatively, the immunoassay may be carried out in liquid phase and a variety of separation methods may be employed to separate the bound labeled component from the unbound labelled components. These methods are known to those of skill in the art and include immunoprecipitation, column chromatography, adsorption, addition of magnetizable particles coated with a binding agent and other similar procedures.

An immunoassay may also be carried out in liquid phase without a separation procedure. Various homogeneous immunoassay methods are now being applied to immunoassays for protein analytes. In these methods, the binding of the binding agent to the analyte causes a change in the signal emitted by the label, so that binding may be measured without separating the bound from the unbound labelled component.

Western blot (immunoblot) analysis can also be used to detect the presence of antibodies to *Chlamydia* in the sample. This technique is a reliable method for confirming the presence of antibodies against a particular protein in the sample. The technique generally comprises separating proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the separated proteins. This causes specific target antibodies present in the sample to bind their respective proteins. Target antibodies are then detected using labeled antihuman antibodies.

The immunoassay formats described above employ labelled assay

components. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide variety of labels may be used. The component may be labelled by any one of several methods. Traditionally a radioactive label incorporating ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P was used. Non-radioactive labels

include ligands which bind to labelled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labelled ligand. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation.

Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labelling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (e.g., streptavidin) molecule which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labelled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Some assay formats do not require the use of labelled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labelled and the presence of the target antibody is detected by simple visual inspection.

Pharmaceutical Compositions

5

10

15

20

25

30

M

The peptides or antibodies (typically monoclonal antibodies) of the present invention and pharmaceutical compositions thereof are useful for administration to mammals, particularly humans, to treat and/or prevent *Chlamydia* infections. Suitable formulations are found in *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985).

10

15

20

25

30

The immunogenic peptides or antibodies of the invention are administered prophylactically or to an individual already suffering from the disease. The peptide compositions are administered to a patient in an amount sufficient to elicit an effective immune response to *Chlamydia*. An effective immune response is one that inhibits infection. An amount adequate to accomplish this is defined as "therapeutically effective dose" or "immunogenically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician, but generally range for the initial immunization (that is for therapeutic or prophylactic administration) from about 0.1 mg to about 1.0 mg per 70 kilogram patient, more commonly from about 0.5 mg to about 0.75 mg per 70 kg of body weight. Boosting dosages are typically from about 0.1 mg to about 0.5 mg of peptide using a boosting regimen over weeks to months depending upon the patient's response and condition. A suitable protocol would include injection at time 0, 4, 2, 6, 10 and 14 weeks, followed by further booster injections at 24 and 28 weeks.

For therapeutic use, administration should begin at the first sign of infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In some circumstances, loading doses followed by boosting doses may be required. The resulting immune response helps to cure or at least partially arrest symptoms and/or complications. Vaccine compositions containing the peptides are administered prophylactically to a patient susceptible to or otherwise at risk of the infection.

The pharmaceutical compositions (containing either peptides or antibodies) are intended for parenteral or oral administration. Preferably, the pharmaceutical compositions are administered parenterally, e.g., subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic polypeptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, e.g., water, buffered water, 0.4% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain

10

. 15

20

25

pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

The compositions may also comprise carriers to enhance the immune response. Useful carriers are well known in the art, and include, e.g., KLH, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly(lysine:glutamic acid), influenza, hepatitis B virus core protein, hepatitis B virus recombinant vaccine and the like.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed ty incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

As noted above, the peptide compositions are intended to induce an immune response to *Chlamydia*. Thus, compositions and methods of administration suitable for maximizing the immune response are preferred. For instance, peptides may be introduced into a host, including humans, linked to a carrier or as a homopolymer or heteropolymer of active peptide units from various *Chlamydia* proteins disclosed here. Alternatively, a "cocktail" of polypeptides can be used. A mixture of more than one polypeptide has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies to a number of epitopes.

The compositions also include an adjuvant. As used here, number of adjuvants are well known to one skilled in the art. Suitable adjuvants include incomplete Freund's adjuvant, alum, aluminum phosphate, aluminum hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP),

N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-Lalanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl

20

25

lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against the immunogenic peptide.

The concentration of immunogenic peptides of the invention in the pharmaceutical formulations can vary widely, i.e. from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

The peptides of the invention can also be expressed by attenuated viral 10 hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (Nature 351:456-460 (1991)). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., Salmonella typhi vectors and the like, will be apparent to those skilled in the art from the description herein.

The DNA encoding one or more of the peptides of the invention can also be administered to the patient. This approach is described, for instance, in Wolff et. al., Science 247: 1465-1468 (1990) as well as U.S. Patent Nos. 5,580,859 and 5,589,466.

In order to enhance serum half-life, the peptides may also be encapsulated, introduced into the lumen of liposomes, prepared as a colloid, or other conventional techniques may be employed which provide an extended serum half-life of the peptides. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., Ann. Rev. Biophys. Bioeng. 9:467 (1980), U.S. Pat. Nos. 4, 235,871, 4,501,728 and 4,837,028.

30

EXAMPLES

The following examples are offered to illustrate, but no to limit the claimed invention.

Example 1:

10

15

20

25

30

This example describes comparison of the *C. pneumoniae* genome disclosed here and the, previously sequenced, *C. trachomatis* genome (Stephens, et al. Science 282:754-759 (1998)).

The apparent low level of DNA homology between *C. trachomatis* and *C. pneumoniae* (Campbell, *et al.*, *J. Clin. Microbiol.* **25**:1911-1916 (1987)) yet analogous cell structures and developmental cycles, predicts that comparative analysis of the two genomes will significantly enhance the understanding of both pathogens. Identification of genes that are present in one species but not the other are of particular importance for the mutually exclusive biological, virulence and pathogenesis capabilities of each.

Identification of genes shared between the two species strongly supports the requirement for these capabilities in a biological system that has, over its long-term association with mammalian host cells, evolved to reduce the metabolic capacities while optimizing survival, growth and transmission of these unique pathogens.

The previously sequenced *C. trachomatis* genome contains 1,042,519 nucleotides and 875 likely protein-coding genes. Similarity searching permitted the inferred functional assignment of sequences 636 (60%) genes disclosed here and 251 (23%) are similar to hypothetical genes for other bacterial organisms including those for *C. trachomatis*. The remaining 186 (17%) genes are not homologous to sequences deposited in GenBank. Seventy *C. trachomatis* genes are not represented in the *C. pneumoniae* genome. These are contained within blocks consisting of 2-17 genes and 19 single genes. Of the 70 *C. trachomatis* genes without homologs in *C. pneumoniae*, 60 are classified as encoding hypothetical proteins. The remaining genes not represented in *C. pneumoniae* consist of the tryptophan operon (*trpA,B,R*), *trpC*, two predicted thiol protease genes, and 4 genes assigned to the phospholipase-D superfamily.

It is evident that there is a high level of functional conservation between C. pneumoniae and C. trachomatis as orthologs to C. trachomatis genes were identified for 859 (80%) of the predicted coding sequences for C. pneumoniae. The level of similarity for individual encoded proteins spans a wide spectrum (22-95% amino acid identity) with an average of 62% amino acid identity between orthologs from the two species. The percent amino acid identity between orthologous chlamydial proteins is similar among functional groups with the highest for proteins associated with translation and the lowest for proteins whose function in chlamydiae is uncharacterized and not related to proteins encoded by other organisms. The gene order of the homologous set of genes in C.

10

15

20

25

30

pneumoniae shows reorganization relative to the genome of *C. trachomatis*; however, there is a high level of synteny for the gene organization of the two genomes. We identified thirty-nine blocks of 2 or more genes whose gene organization is colinear with homologs to *C. trachomatis*, although some of these are inverted. The distribution of genome reorganization is not evenly distributed on the chromosome as the region between *C. pneumoniae* coding sequences 0130-0300 contains substantially more reorganization than other areas of the genome. This region coincides with the predicted chromosome replication terminus.

We identified orthologs of enzymes characterized in other bacteria that account for the essential requirements for DNA replication, repair, transcription and translation including two predicted DNA helicases of the Swi2/Snf2 family found in *C. trachomatis*. Similar to *C. trachomatis*, alternative sigma subunits for RNA polymerase, σ²⁸ and σ⁵⁴, were identified in addition to anti-σ regulatory system factors RsbV, a RsbW-like single-domain histidine kinase, and a RsbU-like protein phosphatase. These findings suggest that the fundamental mechanisms of transcriptional regulation are conserved among *Chlamydia*. The *C. trachomatis* proteins containing SET and SWIB domains, and a SWIB domain fused to the C-terminus of the chlamydial topoisomerase I, not identified outside eukaryotes, are found in *C. pneumoniae* supporting their possible role in the chromatin condensation-decondensation characteristic of the biologically unique chlamydial developmental cycle.

The central metabolic pathways inferred from the *C. pneumoniae* genome sequence are the same as those identified for *C. trachomatis C. pneumoniae* has a glycolytic pathway and a linked tricarboxylic acid cycle, although likely functional, is incomplete as genes for citrate synthase, aconitase, and isocitrate dehydrogenase were not identified. *C. pneumoniae* has a complete glycogen synthesis and degradation system supporting a role for glycogen synthesis and utilization of glucose-derivatives in chlamydial metabolism. Genes encoding essential functions in aerobic respiration are present and electron flux may be supported by pyruvate, succinate, glycerol-3-phosphate, and NADH dehydrogenases, NADH-ubiquinone oxidoreductase and cytochrome oxidase. *C. pneumoniae* also contains the V (vacuolar)-type ATPase operon and the two ATP translocases found in *C. trachomatis*.

The type-III secretion virulence system required for invasion by several pathogenic bacteria and found in the *C. trachomatis* genome in three chromosomal

15

20

25

30

5

10

locationsis also present in the *C. pneumoniae* genome. Each of the components is conserved and their relative genomic contexts are conserved. Genes such as a predicted serine/threonine protein kinase and other genes physically linked to genes encoding structural components of the type-III secretion apparatus, but without identified homologs, are also highly similar between the two species suggesting the functional roles in modifying cellular biology are fundamentally conserved.

Chlamydia-encoded proteins that are not found in chlamydial organisms but localized to the intracellular chlamydial inclusion membrane are likely essential for the unique intracellular biology and perhaps differences in inclusion morphology observed between species of Chlamydia. Several such proteins, termed IncA,B&C, have been characterized for a C. psittaci strain (Rockey, et al. Mol. Microbiol. 15:617-626 (1995); Rockey et al. Infact. Immun. 62:106-112 (1994)). C. pneumoniae and C. trachomatis encode orthologs to C. psittaci IncB and IncC and C. trachomatis also contains an ortholog to IncA. C. pneumoniae contains two genes that encode proteins with similarity to IncA (CPn0186 and CPn0585), although the level of homology is low suggesting analogous but possibily altered functions.

The tryptophan biosynthesis operon (trpA, trpB, trpR) and trpC identified in C. trachomatis is conspicuously missing in the C. pneumoniae genome. This represents the entire repertoire of genes associated with tryptophan biosynthesis identified in C. trachomatis. Seventeen genes adjacent to the C. trachomatis tryptophan operon also were not found in the C. pneumoniae genome. This region is the single largest loss of a contiguous genomic segment and includes 4 HKD superfamily encoding genes that encompass a family of proteins related to endonuclease and phospholipase D. These findings may be important for the ability of Chlamydia to persist in their hosts and cause disease by eliciting potent, focal and persistent inflammatory responses thought to be essential for pathogenesis.

The *C. pneumoniae* genome contains 187,711 additional nucleotides compared to the *C. trachomatis* genome, and the 214 coding sequences not found in *C. trachomatis* account for most of the increased genome size. Eighty-eight of these genes are found in blocks of >10 genes (11-30 genes/block), 41 are single genes, and the remainder are partnered with at least one other gene. Based upon the observation that ~70% of all the *C. pneumoniae* genes have an identifiable homolog in GenBank, exclusive of *C. trachomatis*, it would be expected that over 150 of the 214 genes should

10

15

25

30

have a homolog in GenBank, many associated with a function. However, only 28 coding sequences have similarity to genes from other organisms. Thus the majority of the genes that are mutually exclusive of C. trachomatis (186 of 214), and the 60 of 70 C. trachomatis genes that lacked an identifiable homolog in C. pneumoniae, do not have detectable homologs to genes from other organisms. We predict that most of the unique genes are essential for specific attributes that define the differential biology, tropism and pathogenesis of C. trachomatis and C. pneumoniae. Moreover, this suggests that C. pneumoniae has more unique biological (i.e., virulence) capacity than C. trachomatis. The ability of C. pneumoniae to be more invasive and survive in a broader range of host cell types than C. trachomatis is consistent with this hypothesis. Not all of the differences in biological capacity may be associated with mutually exclusive genes. One explanation for the significantly lower level of homology between protein sequences assigned as having C. pneumoniae and C. trachomatis orthologs but no identifiable orthologs in other organisms is that this set of proteins is not only associated with biological requirements specific for Chlamydia but this polymorphism may account for differential biology between the two species. The determination of the genome sequence from a representative of the C. psittaci group will precisely delineate those genes that are mutually exclusive and specific for each species.

The major functionally identifiable addition to the *C. pneumoniae* genome is a large expansion of genes encoding a new family of chlamydial polymorphic membrane proteins (Pmp), alone representing 22% of the increased coding capacity. While the *C. trachomatis* genome has 9 *pmp* genes, remarkably the *C. pneumoniae* genome contains 21 *pmp* genes. Most of these genes appear to be amplified in two regions of the genome with three stand-alone genes. Interestingly one of the stand-alone genes is most closely related to the *C. trachomatis pmpD* which is the only stand-alone *pmp* gene in the *C. trachomatis* genome and it is located with the same relative genomic context, suggesting an essential and conserved function for this paralog. Six Pmp-coding genes are presumably not functional as five contain predicted coding frame-shifts and one is truncated. The amplification of this gene family and the confidently predicted frame-shifts suggest a specific molecular mechanism to promote functional or antigenic diversity. The biological role of this protein family remains enigmatic, although at least one of the proteins in *C. psittaci* related to this family is exposed on the chlamydial surface.

15

20

25

While a function could not be assigned for most of the unique *C. pneumoniae* genes, several have significant similarity to genes from other organisms. Functional assignments could be made for genes encoding GMP synthetase, IMP dehydrogenase, UMP synthase, uridine kinase, biotin synthase pathway proteins, methylthioadenosine nucleosidase, a DNA glycosylase and aromatic amino acid hydroxylase. Thus a complete pathway was identified for biotin biosynthesis. The additional purine and pyrimidine salvage pathway genes presumably reflect metabolic limitations in one of the cell types that *C. pneumoniae* infects or differences in the ability of *C. pneumoniae* to transport precursor nucleosides or nucleotides.

The addition of aromatic amino acid hydroxylase in *C. pneumoniae* is intriguing especially in light of the loss of tryptophan biosynthetic genes and the inability to synthesize other amino acids including phenylalanine. Aromatic amino acid hyroxlyases include three distinct enzymes that function to receptively oxidize phenylalanine to tyrosine, tyrosine to Dopa, and tryptophan to 5-hydroxytryptophan and serotonin. Although the chlamydial protein is similar to proteins of this family and incrementally more closely related to tryptophan hydroxylase, its specific function could not be confidently predicted. We hypothesize that it may be involved in *C. pneumoniae* virulence. Tryptophan hydroxylase has not been previously identified in bacteria and the origin of the chlamydial gene appears to be from eukaryotes. The functional role of an aromatic amino acid hydroxylase for *C. pneumoniae* is linked to the unique intracellular biology of this organism and may represent a key contribution to *C. pneumoniae* persistence and pathogenesis.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

Ins 30

Table 1 provides functional assignments of *C. pneumoniae* nonproteinencoding genomic sequences. Table 2 provides functional assignments of protein coding sequences. Table 3 provides the amino acid sequences of the proteins corresponding to the coding sequences.

TABLE 1

typ	e	SEQ ID NO:1 start position	SEQ ID NO end positio		ene	
That the main that the same th	Orina PRNA PRNA PRNA PRNA PRNA PRNA PRNA PRNA	start position 841664 138493 607342 1000564 1002415 1005393 269070 164318 296224 836191 1030533 784896 781680 961536 999949 268992 672236 680178 715863 784994 843926 409922 631373 677337 807413 877473 462141 1085605 786780 89728 293477 87522 199301 199390 626904 708359 1142034 1230028	end positio 841396 138074 607649 1002115 1005278 1005509 269142 164389 296151 836119 1030603 784822 781610 961607 1000023 269065 672318 680257 715971 739486 1175944 784922 843999 409848 631445 677264 807341 877400 462214 1085676 786708 89657 293405 87450 199229 199317 626987 708440 1142117 1229945	(R) Putal (R) tmRNA Ribonucle 16S rRNA 23S rRNA 5S rRNA Ala tRNA Asn tRNA (R) Asp t (R) Gly tRNA (R) Gly tRNA His tRNA Leu tRNA (R) Pro tRNA (R) Arg t (R) Thr t (R) Thr t	A ease P RNA _1 ERNA ERNA_2 ERNA ERNA_1 _2 _1 _2 _3 _4 _5 ERNA_1 ERNA_2 ERNA_1 ERNA_2 ERNA_1 ERNA_2 ERNA_1 ERNA_2 ERNA_3 ERNA_1 ERNA_2 ERNA_3 ERNA_1 ERNA_2 ERNA_1 ERNA_3 ERNA_1 ERNA_2 ERNA_1 ERNA_3 ERNA_1 ERNA_2 ERNA_1 ERNA_3 ERNA_1 ERNA_3 ERNA_1 ERNA_3 ERNA_1 ERNA_2 ERNA_1 ERNA_3 ERNA_1 ERNA_3 ERNA_1 ERNA_2 ERNA_1	n of Replica
	ERNA ERNA ERNA ERNA	91070 293399 296147 1137389	90999 293317 296075 1137462	(R) Trp t (R) Tyr t (R) Val t Val tRNA_	RNA RNA_1	· · · · · · · · · · · · · · · · · · ·

TABLE 2

```
Gene Function (C. trachomatis orthogod in parentheses)
                              Strand
Gene 1
                                        CT001 hypothetical protein
                                        gatC-Glu-tRNA Gln Amidotransferase (C subunit) - (CT002)
                                  R
CPROODI
           282
                       875
                                  •
                                        gatA-Glu tRNA Gln Amidotransferae-(CT003)
CPn0002
           573
                                        gat8 (Petil2) Glu FPNA Gln Amidotransferase (8 Subunit) - (CT904)
                                  P
                       2370
           895
CPn0003
                       : (13
                                  Ŀ
                                        pmp_i-Polymorphic Outer Membrane Protein G Family
           2170
 rn0004
                                  F
                       0872
           4127
CPNUUUS
                                  R
                       7141
           7293
CPn0006
                       10496
                                  F
           7605
CPn0007
                       11685
                                  F
           10975
CPn0008
                       13119
           11815
CPn0009
                                  P
                       14325
CPn0010
           13435
                                        frame-shift with 0010
                                  P
                       15746
           14379
CPn0010
                       16614
           15892
CPn0011
                                        pmp_2-Polymorphic Outer Membrane Protein G Family
                       18212
           16644
CPn0012
                                        pmp_3-Polymorphic Outer Membrane Protein G Family
                                  F
                       21106
           18584
CPn0013
                       21922
                                  F
           21392
CPn0014
                                        pmp_3-PMP_3 (frame-shift with 0014)
                       24174
                                  £.
                                        pmp_4-Polymorphic Outer Membrane Protein G Family
           21835
CPn0015
                                  £,
                       26188
            24416
                                        pmp_4-PMP_4 (frame-shift with 0016)
CPn0016
                       27170
                                  F.
                                        pmp_5-Polymorphic Outer Membrane Protein G Family
           26094
CPn0017
                                  F
                       29003
            27522
                                        pmp_5-PMP_5 (frame-shift with 0018)
CPn0018
                                        Predicted OMP [leader (14) peptide: outer membrane]-(CT351)
                       30356
                                  F
           29007
CP00019
                       30603
                                        Predicted OMP [leader (19) peptide]-(CT350)
CPn.0020
            32687
                       32707
            34410
CP00021
                                        maf-(CT349)
                       34395
                                        yjjK/alr-ABC Transporter Protein ATPase-(CT348)
            34982
CPhi0022
                       35014
            36603
CPm0023
                                        xerC-Integrase/recombinase-(CT347)
                                        elaC/atsA-Sulphohydrolase/Glycosulfatase-( CT346)
                       36661
CPn0024
            37596
                       37684
                                   R
            38604
                                        CT345 hypothetical protein-(CT345)
CP50025
                                   R
                       38762
            39625
CPh0026
                                        lon-Lon ATP-dependent Protease-(CT344)
                                   R
                       39778
CP00027
            42234
                                   R
                       42543
            43325
CPp0028
            43755
                       43390
                                   R
                                        gcp_1-0-Sialoglycoprotein Endopeptidase_1-(CT343)
CPn0029
                                   P
                       44529
CPH0030
            43891
                                         rs21-S21 Ribosomal Protein-(CT342)
                       44884
            44711
CP#0031
                                         dnaJ-Heat Shock Protein J-(CT341)
                                        pdhA&B/odbA&odbB-(pyruvate) Oxoisovalerate Dehydrogenase Alpha & Beta
                                   P
                       46098
 CPh0032
            44923
                       48171
                                   F
            46138
 CP#@033
                                           Fusion-(CT340)
 CPn0034
                        48210
                                   R
            49457
                                         CT339 hypothetical protein
                       49569
                                   R
 CPh0035
            51029
                                         CT338 hypothetical protein
                                   F
                        51796
                                         ptsH-PTS Phosphocarrier Protein Hpr-(CT337)
 CPn0036
            51002
                                   F
            51792
                        52115
 CPn0037
                                         ptsI-PTS PEP Phosphotransferase-(CT336)
                                   F
                        53831
            52119
 CPn0038
                                         ybaB-(CT335)
                        53963
                                   R
            54250
 EPn0039
                                         dnax_1-DNA Pol III Gamma and Tau_1-(CT334)
                        54318
            55643
 CPn0040
            55996
                        57342
                                   F
 CPn0041
                        58182
                                   F
 CPn0042
            57403
            58447
                        60372
                                   F
 CPn0043
                        60778
            60419
 CPn0044
                                   F
            61069
                        62790
 CPn0045
                        63263
                                   F
            62790
 CPn0046
                                         *YqfF-Bs conserved hypothetical IM protein
                                   F
 CPn0047
            63455
                        63652
                        65801
            63687
 CPn0048
                                   R
             66296
                        65817
 CPn0049
             66813
                        66499
                                   R
 CPn0050
                                   F
             66833
                        67111
 CPn0051
                                         hemC-Porphobilinogen Deaminase-(CT299)
            68005
                        67304
---CPn0052-
                                         sms-Sms Protein-(CT298)
             69344
                        67986
                                   R
 CPn0053
                                         rnc-Ribonuclease III-(CT297)
             70023
                        69313
                                    R
 CPn0054
                                         CT296 hypothetical protein
                        70590
             70129
                                   F
 CPn0055
                                         mrsA-Phosphomannomutase-(CT295)
             70953
                        72746
                                         sodM-Superoxide Dismutase (Mn)-(CT294)
 CPn0056
                                         accD-AcCoA Carboxylase/Transferase Beta-(CT293)
                        73554
             72934
 CPn0057
                        74562
             73639
 CPn0058
                                         dut-dUTP Nucleotidohydrolase-(CT292)
                        75050
             14616
 CPn0059
                                         ptsN_1-PTS IIA Protein-(CT291)
                                         ptsN_2-PTS IIA Protein + HTH DNA-Binding Domain-(CT290)
                        75528
 CPn0060
             75055
                        76208
                                    F
             75534
 CPROOFL
                                         CT289 hypothetical protein
                        77690
             76308
  CPn0062
                        78267
  C2n0063
             78112
                         78576
             78346
  CPR0054
                                         CT288 hypothetical procein
                         8065 L
                                    F
 CPn0055
             78924
                         82655
             40925
  CPROOSS
```

```
84053
             82953
  CPn0067
                                          CT360 hypothetical protein
             84903
                         84331
                                    R
  CPn0068
                         87086
  CPn0069
             85236
                         87208
                                    R
             87378
  CPn0070
                                          CT325 hypothetical protein
                         87599
             88045
  CPn0071
                                          CT324 hypothetical protein
                         88057
                                    R
             89061
  CPn0072
                                          infA-Initiation Factor IF-1-(CT323)
                                    F
                         89574
             89356
  CPn0073
                                          tufA-Elongation Factor Tu-(CT322)
                         90955
                                    F
             89774
  CPn0074
                                          secE-preprotein translocase-(CT321)
                         91350
             91102
  CPn0075
                                          nusG-Transcriptional Antitermination-(CT320)
                         91903
             91358
  CPn0076
                                          rlll-Lll Ribosomal Protein-(CT319)
                                    F
             92013
                         92435
  CPn0077
                                          rll-Ll Ribosomal Protein-(CT318)
                         93160
             92465
  CPn0078
                                          rll0-Ll0 Ribosomal Protein-(CT317)
                         93688
  CPn0079
             93179
                                          rl7-L7/L12 Ribosomal Protein-(CT316)
                         94121
             93735
  CPn0080
                                          rpoB-RNA Polymerase Beta-(CT315)
             94261
                         98016
                                    F
  CPn0081
                                          rpoC-RNA Polymerase Beta' -(CT314)
                         102221
             98043
  CPn0082
                                          tal-Transaldolase-(CT313)
             102332
                                    F
                         103312
  CPn0083
                                          predicted ferredoxin-(CT312)
                         103751
             103362
  CPn0084
                                          CT311 hypothetical protein
                         103766
                                    R
  CPn0085
             104506
                                          atpE-ATP Synthase Subunit E-(CT310)
                         105527
                                     F
             104904
  CPn0086
                                          CT309 hypothetical protein
                                    F
                         106376
  CPn0087
             105579
                                          atpA-ATP Synthase Subunit A-(CT308)
                         108145
  CPn0088
             106373
                                          atpB-ATP Synthase Subunit B-(CT307)
                                    F
                         109466
             108153
  CPn0089
                                          atpD-ATP Synthase Subunit D-(CT306)
                         110080
                                     F
  CPn0090
             109454
                                          atpI-ATP Synthase Subunit I-(CT305)
                                    F
                         112053
  CP50091
              110074
                                          atpK-ATP Synthase Subunit K-(CT304)
                         112573
                                     F
  CPN0092
              112151
                                          CT303 hypothetical protein
                                     F
                         113015
             112509
  CPn0093
                                          valS-Valyl tRNA Synthetase-(CT302)
                         115971
                                     F
              113152
   CP50094
                                          pknD-S/T Protein Kinase-(CT301)
   CP0095
                         118790
                                     F
              116037
                                          uvrA-Excinuclease ABC Subunit A-(CT333)
                         118837
                                     R
   CÉn 0096
              124314
                                          pyk-Pyruvate Kinase-(CT332)
   CPh0097
              124555
                         126006
                                     F
                                          htrB-Acyltransferase-(CT010)
                         126091
                                     R
   CPh0098
              127491
                         127865
                                     F
              127593
   CPn0099
                                          CT011 hypothetical protein
                         127882
                                     R
   CPn0100
              129141
                                          ybbP family hypothetical protein-(CT012)
                         129141
                                     R
   CPn0101
              129932
                                          cydA-Cytochrome Oxidase Subunit I-(CT013)
                         131466
                                     F
   CPR0102
              130123
                                          cydB-Cytochrome Oxidase Subunit II-(CT014)
                         132511
                                     F
              131480
   C₽#0103
                                          CT017 hypothetical protein
                         132676
                                     R
   CPp0104
              133875
                                          CT016 hypothetical protein
                         134029
                                     R
              134847
   CPa:0105
                                          phoH-ATPase-(CT015)
                         136374
                                     F
   CP00106
              135091
                                          CT058 hypothetical protein_1
                          136392
                                     R
   CPD0107
              137162
                         137303
                                          CT018
                                     R
   CPn0108
              137857
                                           ileS-Isoleucyl-tRNA Synthetase-(CT019)
              138655
                          141783
                                     F
   CPn0109
                                           lepB-Signal Peptidase I-(CT020)
                                     R
                         141827
   CPn0110
              143734
                                           CT021 hypothetical protein
              144686
                          143934
                                     R
   CPn0111
                                           rl31-L31 Ribosomal Protein-(CT022)
                         145093
                                     F
   CPn0112
              144767
                                           pfrA-Peptide Chain Releasing Factor (RF-1)-(CT023)
   CPn0113
              145335
                          146405
                                     F
                                           hemK-A/G specific methylase-(CT024)
                                     F
                          147261
   CPn0114
              146398
                                           ffh-Signal Recognition Particle GTPase-(CT025)
                          148622
                                     F
   CPn0115
              147279
                                           rs16-S16 Ribosomal Protein-(CT026)
                          148972
                                     F
   CPn0116
              148616
                                           trmD-tRNA (guanine N-1)-Methyltransferase-(CT027)
                          150071
                                     F
   CPn0117
              148989
                                           rl19-L19 Ribosomal Protein-(CT028)
                          150464
                                     F
              150102
   CPn0118
                                           rnhB_1-Ribonuclease HII_1-(CT029)
   CPn0119
              150523
                          151164
                                     F
                                           gmk-GMP Kinase-(CT030)
                          151778
   CPn0120
              151164
                                           CT031 hypothetical protein
                          152068
   CPn0121
              151778
                                           metG-Methionyl-tRNA Synthetase-(CT032)
              152071
                          153723
                                     F
                                           recD_1-Exodeoxyribonuclease V (Alpha Subunit)_1-(CT033)
   CPn0122
                          153774
___CPn0123_
              _155969_
                          158068
                                     F
   CPn0124
              156614
   CPn0125
              158096
                          158605
                                     F
                                     F
              158809
                          161085
   CPn0126
                                           ytfF-Cationic Amino Acid Transporter-(CT034)
   CPn0127
               162143
                          161130
                                      R
                                           bpl1-Biotin Protein Ligase-(CT035)
                                     F
   CPn0128
              162277
                          163053
                                           similarity to CT036
   CPn0129
               163717
                          163064
                                      R
   CPn0130
               164245
                          163751
                                      R
                          165580
                                      F
   CPn0131
               164549
   CPn0132
               165587
                          166561
                                      F
                                           CHLPS hypothetical protein-(CT109)
                          166564
                                      R
   CPn0133
               167334
                                           groEL_1-HSP-60_1-(CT110)
   CPn0134
               169098
                          167467
                                      R
                                           groES-10KDa Chaperonin-(CT111)
                          169143
                                      R
   CPn0135
               169448
                                           pepF-Oligopeptidase-(CT112)
   CPn0136
               171401
                          169569
                                      R
                                           ybgI-ACR family-(CT108)
                          171502
                                      R
   CPn0137
               172254
                                           hemL-Glutamate-1-semialdehyde-2,1-aminomutase-(CT210)
                          172700
   CPn0138
               174019
```

```
- CPn0139
            174656
                        174093
                                   R
                                         yqgE-(CT210)
                                   R
                                         yqdE-(CT212)
            175110
                        174673
 CPn0140
                                         rpiA-Ribose-5-P Isomerase A-(CT213)
                        175110
                                   R
 CPn0141
            175802
                        175816
                                   R
            176091
 CPn0142
                                         *yxjG_Bs_1 Hypothetical Protein
                        176214
                                   R
 CPn0143
            177335
                                         clpB-Clp Protease ATPase-(CT113)
                        180560
                                   F
            177963
 CPn0144
                                         CT114 hypothetical protein
                        182369
            180777
 CPn0145
                        183095
 CPn0146
            182613
            183225
                        183671
CPn0147
                                         pknl-S/T Protein Kinase-(CT145)
                        185702
                                   F
            183846
 CPn0148
                        187700
                                         dnlJ-DNA Ligase-(CT146)
 CPn0149
            185715
                                         CT147 hypothetical protein
                                   F
            187834
                        192444
 CPn0150
                                         mhpA-Monooxygenase-(CT148)
                        192625
CPn0151
            194142
                                         CT149 hypothetical protein
            195265
                        194318
                                   R
CPn0152
                                         leuS-Leucyl tRNA Synthetase-(CT209)
            195433
                        197892
CPn0153
                                         gseA-KDO Transferase-(CT208)
 CPn0154
            197892
                        199202
                                   F
            199691
                        199488
CPn0155
 CPn0156
            200117
                        199770
                                   R
            200723
                        200298
                                   R
 CPn0157
 CPn0158
            201430
                        200894
                                   R
            201772
                        201467
 CPn0159
                                   R
                                         pfkA_1-Fructose-6-P Phosphotransferase_1-(CT207)
 CPn0160
            203791
                        202127
                                   R
                        203798
                                   R
                                         predicted acyltransferase family-(CT206)
            204622
 CPn:0461
 CPn0162
            205828
                        204803
                                   R
 CPn 0163
                        206394
            206026
            206498
                        206998
CPn0164
                                   P
            206998
                        207582
CPn 0165
                        207962
СЪщ63 66
            207630
CPn0167
                        207977
            208306
                                   R
CPn 0168
            208641
                        208417
                        208710
CP:10169
            209501
                                   R
CPn 0170
            211026
                        210025
                                   R
                                         *quaA-GMP Synthase
CPn0171
            212435
                        211149
                                   R
                                         *guaB/impD-Inosine 5'-monophosphase dehydrogenase (C00H-terminal region
            213177
                        212440
CPn0172
CPn 0173
            213987
                        213715
                                   R
CPm0174
            214257
                        214724
                                   F
CPn0175
            214898
                        215275
                                         CT153 hypothetical protein
CPn 0176
            215286
                        216518
                                   F
CPri 0 177
            217459
                        216608
                                   R
CPn 0178
            218052
                        217789
                                   R
CPn0179
            218403
                        218056
CPn0180
            218851
                        218355
                                   R
CPn0181
            219175
                        218777
                                   R
                                         accC-Biotin Carboxylase-(CT124)
CPn0182
            220695
                        219334
                                   R
CPn0183
            221195
                        220695
                                         accB-Biotin Carboxyl Carrier Protein-(CT123)
                                   R
                                         efp_1-Elongation Factor P_1-(CT122)
CPn0184
            221775
                        221221
                                   R
CPn0185
            222451
                        221765
                                   R
                                         rpe/araD-Ribulose-P Epimerase-(CT121)
                                         *similarity to Cps IncA_1-(CT119)
CPn0186
            222899
                        224068
                                   F
CPn0187
            224248
                        225045
                                         predicted methylase-(CT133)
                                   F
CPn0188
            225111
                        226400
                                   F
                                         CT132 hypothetical protein
                                         CT131 homolog-(Possible Transmembrane Protein)
CPn0189
            226400
                        229825
                                   F
CPn0190
            229919
                        231274
                                         glnQ-ABC Amino Acid Transporter ATPase-(CT130)
CPn0191
            231991
                        231314
                                   R
                                         glnP-ABC Amino Acid Transporter Permease-(CT129)
CPn0192
            232634
                        231984
                                         *argR-Arginine Repressor
CPn0193
            233126
                        232686
                                   R
                                        gcp_2-O-Sialoglycoprotein Endopeptidase_2-(CT197)
CPn0194_
                        234241
            .233210...
                                         oppA_1-Oligopeptide Binding Protein_1
                        235785
CPn0195
            234190
                                         oppA_2-Oligopeptide Binding Protein_2-(CT198)
CPn0196
            235939
                        237519
                                         oppA_3-Oligopeptide Binding Protein_3
CPn0197
                        238882
            237578
                                   F
                                         oppA_4-Oligopeptide Binding Protein_4
CPn0198
            239169
                        240746
                                         oppB_1-Oligopeptide Permease_1-(CT199)
CPn0199
            241042
                        241983
                                   F
                                         oppC_1-Oligopeptide Permease_1-(CT200)
CPn0200
            242017
                        242868
                                         oppD-Oligopeptide Transport ATPase-(CT201)
CPn0201
            242864
                        243715
CPn0202
                        244500
                                   F
                                         oppF-Oligopeptide Transport ATPase-(CT202)
            243715
CPn0203
            245008
                        245802
                                   F
CPn0204
            245817
                        246002
CPn0205
            246133
                        246327
CPn0206
                        247161
                                   F
                                         CT203 hypothetical protein
            246409
                                         ybhI/sodiT1-Oxoglutarate/Malate Translocator-(CT204)
CPn0207
            247208
                        248617
                                         pfkA_2-Fructose-6-P Phosphotransferase_2-(CT205)
CPn0208
                        250602
            248953
CPn0209
            251036
                        251272
```

```
252384
 CPn0210
                        251440
                                   R
            252756
 CPn0211
                        252463
            254066
                        252888
 CPn0212
                                   R
 CPn0213
            254342
                        254190
                                   R
            255657
                        254446
 CPn0214
                                   R
 CPn0215
            257015
                        255759
 CPn0216
            257608
                        257174
                                   R
            257896
                        258579
                                   F
                                         ypdP-(CT140)
 CPn0217
 CPn0218
            259058
                        258582
                                   R
            259357
                        260472
                                         tgt-Queuine tRNA Ribosyl Transferase-(CT193)
                                   F
 CPn0219
 CPn0220
            260696
                        261238
                                   F
            261657
                        262064
                                   F
 CPn0221
 CPn0222
            262504
                        262842
                                   F
                                         *weak similarity to Bacteriophage CHP1 (Orf4)
            262956
                        263333
 CPn0223
                                   F
 CPn0224
            263435
                        263674
                                   F
 CPn0225
            263873
                        264541
                                   F
 CPn0226
            264566
                        264967
                                   F
 CPn0227
            265416
                        265009
                                         dsbB-Disulfide bond Oxidoreductase-(CT176)
                                   R
 CPn0228
            266110
                        265412
                                   R
                                         dsbG-Disulfide Bond Chaperone-(CT177)
                        267560
                                         CT178 hypothetical protein
 CPn0229
            266328
                                   F
 CPn0230
            268253
                        267576
                                   R
                                         CT179 hypothetical protein
                                         tauB-ABC Transport ATPase (Nitrate/Fe)-(CT180)
 CPn0231
            268957
                        268253
                                   R
                                         *similarity to 5'-Methylthioadenosine / S-Adenosylhomocysteine
 CPM0232
            270122
                        269232
                                   R
                                            Nucleosidase
 CPn0233
            270424
                        270248
                                   R
            271240
                                         CT181 hypothetical protein
 CPn0234
                       270548
                                   R
 CPn0235
            271416
                       272177
                                         kdsB-deoxyoctulonosic Acid Synthetase-(CT182)
                                   F
            272156
 CPn0236
                       273766
                                   F
                                         pyrG-CTP Synthetase-(CT183)
                                         yggF Family-(CT184)
CPH0237
            273762
                       274214
                                   F
CP10238
            274303
                       275838
                                   F
                                         zwf-Glucose-6-P Dehyrogenase-(CT185)
 CPn0239
            275899
                       276672
                                         devB-Glucose-6-P Dehyrogenase (DevB family)-(CT186)
CPn0240
            277861
                       276698
                                   R
CPn0241
            279354
                       278203
                                   R
CP#0242
            279918
                       279487
                                   R
CPm0243
            280555
                       280133
                                   R
CPn0244
            280918
                       281556
                                         adk-Adenylate Kinase-(CT128)
CPn0245
            281645
                       282499
                                         ydhO-Polysaccharide Hydrolase-Invasin Repeat Family-(CT127)
                                   F
CPn0246
            282952
                       282551
                                   R
                                         rs9-S9 Ribosomal Protein-(CT126)
CP00247
            283415
                       282969
                                        rll3-Ll3 Ribosomal Protein-(CT125)
                                  R
CPn0248
            284327
                       283650
                                   R
                                        ycfV/ybbA-ABC Transporter ATPase-(CT152)
CPn0249
            285841
                       284333
                                  R
                                        CT151 hypothetical protein
CPn0250
            286057
                       285902
                                  R
                                        rl33-L33 Ribosomal Protein-(CT150)
CPn0251
            286060
                       287559
                                  F
                                         *conserved hypothetical protein
CPn0252
            288112
                       287576
                                  R
                                        CT144 hypothetical protein (frame-shift with 0253?)
CPn0253
            288456
                       287950
                                  R
                                        CT144 hypothetical protein_1
CPn0254
            289262
                       288459
                                  R
                                        CT143 hypothetical protein_1
CPn0255
           290165
                       289329
                                  R
                                        CT142 hypothetical protein_1
CPn0256
           291264
                       290398
                                  R
                                        CT144 hypothetical protein_2
CPn0257
           292127
                       291267
                                  R
                                        CT143 hypothetical protein_2
CPn0258
           292534
                                        CT142 hypothetical protein (frame-shift with 0259?)
                       292133
                                  R
CPn0259
           292986
                       292441
                                  R
                                        CT142 hypothetical protein_2
CPn0260
           294045
                       293548
                                  R
                                        secA 1-Protein Translocase Subunit_1-(CT141)
CPn0261
           294302
                       295033
                                  F
                                        ydaO-PP-Loop Superfamily ATPase-(CT217)
CPn0262
           295091
                       295933
                                  F
                                        surE-SurE-like Acid Phosphatase-(CT218)
CPn0263
           296249
                       297136
                                  F
                                        yqfU hypothetical protein-(CT221)
CPn0264
           297730
                       297155
                                  R
                                        ubiD-Phenylacrylate Decarboxylase-(CT220)
CPn0265
           298620
                       297730
                                        ubiA_Benzoate_Octaphenyltransferase-(CT219)
CPn0266
           299184
                       299876
                                  F
CPn0267
           300122
                       300910
                                  F
CPn0268
           300935
                       301318
                                  F
CPn0269
           302450
                       301476
                                        Dipeptidase-(CT138)
                                  R
CPn0270
           303325
                       302468
                                        ywlC-SuA5 Superfamily-related Protein-(CT137)
                                  R
CPn0271
           303634
                       304362
                                  F
                                        Lysophospholipase esterase-(CT136)
CPn0272
           305233
                       304340
                                  R
                                        dnaX_2-DNA Pol III Gamma and Tau_2-(CT187)
CPn0273
           305844
                       305227
                                  R
                                        tdk-Thymidylate Kinase-(CT188)
CPn0274
                       305852
           308353
                                  R
                                        gyrA_1-DNA Gyrase Subunit A_1-(CT189)
CPn0275
           310786
                       308372
                                  R
                                        gyrB_1-DNA Gyrase Subunit B_1-(CT190)
CPn0276
                                        CT191 hypothetical protein
           311137
                       310793
                                  R
CPn0277
           311910
                       311404
                                  R
CPn0278
           312875
                       312060
                                        *conserved outer membrane lipoprotein protein
CPn0279
           313537
                       312875
                                        *Possible ABC Transporter Permease Protein
                                  R
CPn0280
           314572
                       313550
                                        dppF-Dipeptide Transporter ATPase-(CT689)
```

```
dhnA-Predicted 1,6-Fructose Biphosphale Aldolase (dehydrin family) -
 CPn0281
            315057
                       316103
                                        xasA/gadC-Amino Acid Transporter-(CT216)
            316126
                       317529
                                  F
 CPn0282
            318497
                       317532
 CPn0283
            319045
                       318551
 CPn0284
            320595
                       319051
                                  R
 CPn0285
                                        mgtE-Mg++ Transporter (CBS Domain)-(CT194)
 CPn0286
            322059
                       320650
                       322089
 CPn0287
            324221
                                        CT195 hypothetical protein
 CPn0288
            325716
                       324571
                                        aaaT-Neutral Amino Acid (Glutamate) Transporter-(CT230)
                       326996
                                  F
            325812
 CPn0289
                                        Na-dependent Transporter-(CT231)
            327042
                       328523
 CPn0290
                                        incB-Inclusion Membrane Protein B-(CT232)
            328667
                       329194
                                  F
 CPn0291
                                        incC-Inclusion Membrane Protein C-(CT233)
            329228
                       329836
 CPn0292
                                        CT234 hypothetical protein
            329949
                       332723
                                  F
 CPn0293
                                        cAMP-Dependent Protein Kinase Regulatory Subunit-(CT235)
            333092
                       333502
 CPn0294
                                        acpP-Acyl Carrier Protein-(CT236)
 CPn0295
            333863
                       333627
                                        fabG-Oxoacyl (Carrier Protein) Reductase-(CT237)
            334765
                       334022
 CPn0296
                                        fabD-Malonyl Acyl Carrier Transcyclase-(CT238)
                       334774
 CPn0297
            335697
                                  :3
                                        fabH-Oxoacyl Carrier Protein Synthase III-(CT239)
            336721
                       335717
                                  R
 CPn0298
                                        recR-Recombination Protein-(CT240)
 CPn0299
            336816
                       337415
                                  17
                                        yaeT-Omp85 Analog-(CT241)
            337783
                       340152
 CPn0300
                                        (OmpH-Like Outer Membrane Protein) - (CT242)
 CPn0301
            340250
                       340762
                                  1:
                                        lpxD-UDP Glucosamine N-Acyltransferase-(CT243)
            340787
                       341866
                                  ŧ.
 dPh0302
                                        CT244 hypothetical protein
 <u>c⊉n</u>0303
            342958
                       341921
                                  7?
                                        pdhA/odpA-Pyruvate Dehydrogenase Alpha-(CT245)
            343133
                       344158
                                  F
 CPn0304
                                        pdhB/odpB-Pyruvate Dehydrogenase Beta-(CT246)
 ÇPn0305
            344154
                       345137
                                  Ι
                                        pdhC-Dihydrolipoamide Acetyltransferase-(CT247)
            345145
                       346431
 de 0306
                                  1
                                        glgP-Glycogen Phosphorylase-(CT248)
 CPn0307
            348986
                       346515
                                  \mathbf{F}_{i}
                                        similarity to CT249
            349234
                       349596
                                  F.
 @Pn0308
                                        dnaA_1-Replication Initiation Protein_1-(CT250)
 ÇPn0309
            350974
                       349595
                                  R
                                        60IM-60kDa Inner Membrane Protein-(CT251)
 ÇPn0310
                       351049
          . 353433
                                  R
                                        lgt-Prolipoprotein Diacylglycerol Transferase-(CT252)
 dPn0311
            354438
                       353575
                                  R
            354524
                       354976
                                  F
                                        CT101 hypothetical protein
 @Pn0312
                                        acpS-Acyl-carrier Protein Synthase-(CT100)
 €2n0313
            354990
                       355355
                                        trxB-Thioredoxin Reductase-(CT099)
            356285
                       355353
                                  R
 <u>©Pn</u>0314
                                        rs1-S1 Ribosomal Protein-(CT098)
            356977
                       358716
                                  F
  Pn0315
                                        nusA-N Utilization Protein A-(CT097)
 ĠPn0316
            358820
                       360121
                                  F
                                        infB-Initiation Factor-2-(CT096)
                       362750
 ₫₽n0317
            360081
                       363126
                                  F
                                        rbfA-Ribosome Binding Factor A-(CT095)
            362767
 CPn0318
                                        truB-tRNA Pseudouridine Synthase-(CT094)
 CPn0319
            363175
                       363879
                                  F
 Ĉħ0320
                                        ribF-FAD Synthase-(CT093)
            363860
                       364783
                                  F
                                       ychF-GTP Binding Protein-(CT092)
                       364767
 CPn0321
            365858
                                        yscu-YopS Translocation Protein U - (CT091)
            366249
 CPn0322
                       367328
                                  F
                                        lcrD- Low Calcium Response D-(CT090)
 CPn0323
            367331
                       369460
                                  F
                                        lcrE- Low Calcium Response E-(CT089)
 CPn0324
            369492
                       370688
                                  F
                                        sycE-Secretion Chaperone-(CT088)
 CPn0325
            370708
                       371148
                                       malQ-Glucanotransferase-(CT087)
 CPn0326
                       372725
            371148
                                  F
                                        rl28-L28 Ribosomal Protein-(CT086)
 CPn0327
            372945
                       373211
                                  F
                                        CT085 hypothetical protein
 CPn0328
            373241
                       374992
                                  F
                                        Phopholipase D Superfamily [leader (33) peptide]-(CT084)
 CPn0329
            375088
                       376146
                                  F
                                       CT083 hypothetical protein
 CPn0330
            376675
                       376202
                                  R
                                        CT082 hypothetical protein
 CPn0331
            378437
                       376701
                                  R
                                       CHLTR T2 Protein-(CT081)
 CPn0332
            378655
                       378536
                                  R
 CPn0333
            379090
                       378800
                                       ltuB-(CT080)
                                  Ř
 CPn0334
            379311
                       379823
                                        CT079 similarity
                                  F
 CPn0335
                                        folD-Methylene Tetrahydrofolate Dehydrogenase-(CT078)
            379817
                       380674
                                  F
                                       _CPn0336 .__380650 _
                       381591
                                 - F-
 CPn0337
            382027
                       381575
                                        smpB- Small Protein B-(CT076)
                                  R
                                        dnaN-DNA Pol III (beta chain)-(CT075)
CPn0338
            382278
                       383375
                                  F
                                        recF-ABC superfamily ATPase-(CT074)
 CPn0339
            383420
                       384034
                                  F
                                        (frame-shift with 0339)
CPn0340
            383842
                       384156
                                  F
                                        (frame-shift with 0340)
 CPn0341
            384160
                       384495
                                  F
                                        predicted OMP [leader (19) peptide]-(CT073)
CPn0342
            384622
                       385062
                                  F
 CPn0343
            384999
                                  F
                                        (frame-shift with 0342?)
                       385595
                                       yaeL-Metalloprotease-(CT072)
CPn0344
            387420
                       385558
                                  R
CPn0345
            388572
                                       yaeM-(CT071)
                       387436
                                  R
CPn0346
                                        troD/ytgD-Integral Membrane Protein-(CT070)
            389675
                       388704
                                  R
                                        troC/ytgC-Integral Membrane Protein-(CT069)
 CPn0347
            391021
                       389678
                                  R
CPn0348
                                        troB/ytgB-ABC transporter ATPase-(CT068)
            391803
                       391027
                                  R
                                        troA/ytgA-Solute Protein Binding Family-(CT067)
CPn0349
            392770
                       391790
                                  R
CPn0350
                                        CT066 hypothetical protein
            393181
                       393684
                                        adt_1-ADP/ATP Translocase_1-(CT065)
CPn0351
            393888
                       395432
```

```
CPn0352
           395574
                       396830
                                   F
           396893
                       397135
                                   F
CPn0353
           397167
                       398507
                                  F
CPn0354
           399889
                       398591
CPn0355
                                  R
           400459
                       400109
CPn0356
           401317
                       400469
                                  R
CPn0357 .
CPn0358
           401751
                       401578
                                  R
                       403817
                                        lepA-GTPase-(CT064)
           402012
                                  F
CPn0359
                                        gnd-6-Phosphogluconate Dehydrogenase-(CT063)
CPn0360
           405358
                       403922
                                  R
                       405382
                                        tyrs-tyrosyl tRNA Synthetase-(CT062)
           406647
                                  R
CPn0361
                                        fliA/rpsD-Sigma-28/WhiG Family-(CT061)
GPn0362
           407825
                       407055
                                  R
                                        flhA-Flagellar Secretion Protein-(CT060)
                       407943
           409688
CPn0363
                                  R
                       410238
                                        fer4-Ferredoxin IV-(CT059)
CPn0364
           409966
                       411544
CPn0365
           410528
                                  F
CPn0366
           411976
                       412440
                                  F
                       413836
                                  F
CPn0367
           413102
CPn0368
           413790
                       414107
                       415562
                                        CT058 hypothetical protein_2
CPn0369
                                  F
           414351
                                        CT058 hypothetical protein_3
CPn0370
           415800
                       416912
                                  F
                       417503
CPn0371
           417147
                                  F
CPn0372
           417687
                       418001
                                        gcpE-(CT057)
CP<u>11</u>0373
           418380
                       420218
                                  F
CPn0374
           420218
                       420961
                                  F
                                        CT056 hypothetical protein
CPn0375
           421121
                       421615
                                  F
           421854
                       422294
                                  F
CPn0376
                                        sucB_1-Dihydrolipoamide Succinyltransferase_1-(CT055)
CPn 0377
           423438
                       422347
                                        sucA-Oxoglutarate Dehydrogenase-(CT054)
           426168
                       423445
CPT0378
                                  R
                                        CT053 hypothetical protein
CP70379
           426322
                       426765
                                        hemN_1-Coproporphyrinogen III Oxidase_1-(CT052)
           426758
                       427876
CP00380
CPR0381
           429809
                       428037
                                        CT326 similarity
                                        yabC/yraL-SAM-Dependent Methytransferase-(CT048)
CPn0382
           430749
                       430036
                                  R
CPn0383
           431693
                       430749
                                  R
                                        CT047 hypothetical protein
                                        hctB-Histone-like Protein 2-(CT046)
CPT0384
           432377
                       431862
                                  Ŕ
                                        pepA-Leucyl Aminopeptidase A-(CT045)
CPn 0385
           434018
                       432522
CPn0386
           434525
                                        ssb-SS DNA Binding Protein-(CT044)
                       434046
                                  R
CPn0387
           435196
                       434699
                                        CT043 hypothetical protein
                                  R
                                        glgX-Glycogen Hydrolase (debranching)-(CT042)
           435329
                       437320
CPn0388
                                  F
CPn0389
           438134
                       437319
                                        CT041 hypothetical protein
                                  R`
CP0390
           439144
                       438134
                                        ruvB-Holliday Junction Helicase-(CT040)
                                  R
CPn0391
           439692
                       439510
                                  R
CPn0392
           439814
                       440383
                                        dcd-dCTP Deaminase-(CT039)
                                  F
CPn0393
           440379
                       440723
                                        CT038 hypothetical protein
                                        tlyC_1-CBS Domain protein (Hemolysin Homolog)_1-(CT256)
CPn0394
           440736
                     441968
                                  F
                                        CT257 hypothetical protein
CPn0395
           441964
                       443175
                                       yhfo-NifS-related protein-(CT258)
CPn0396
           444353
                       443241
                                  R
CPn0397
           445115
                       444381
                                        PP2C phosphatase family-(CT259)
CPn0398
           445533
                       445700
                                  F
CPn0399
                                       CT253 hypothetical protein
           445879
                       446523
                                  F
CPn0400
           446536
                       447306
                                  F
                                       CT254 hypothetical protein
CPn0401
           447884
                       447495
                                  R
                                       CT255 hypothetical protein
CPn0402
                                       mutY-Adenine Glycosylase-(CT107)
           448994
                      447888
                                  R
CPn0403
           449015
                       449710
                                  F
                                       yceC-predicted pseudouridine synthetase family-(CT106)
CPn0404
           450887
                       449871
                                  R
CPn0405
           451739
                       450966
                                        CT105 hypothetical protein
                                  R
                                        fabI-Enoyl-Acyl-Carrier Protein Reductase-(CT104)
CPn0406
           451969
                      452865
                                  F
CPn0407
                                       HAD superfamily hydrolase/phosphatase-(CT103)
           453742
                       452858
                                  R
                                       CT102 hypothetical protein
CPn04.08
           454105
                      454581
                                 F
CPn0409
           454645
                       455127
                                       CT260 hypothetical protein
                                  F
CPn0410
                                        dnaQ_1-DNA Pol III Epsilon Chain_1-(CT261)
           455123
                       455833
CPn0411
           455833
                       456609
                                  F
                                       CT262 hypothetical protein
CPn0412
           456590
                       457246
                                       CT263 hypothetical protein
                                  F
CPn0413
           459203
                       457227
                                       msbA-Transport ATP Binding Protein-(CT264)
                                  R
                                        accA-AcCoA Carboxylase/Transferase Alpha-(CT265)
CPn0414
           460143
                       459172
CPn0415
           461498
                      460221
                                       CT266 hypothetical protein
                                  R
CPn0416
           461856
                       461557
                                       himD/ihfA-Integration Host Factor Alpha-(CT267)
                                  R
                                       amiA-N-Acetylmuramoyl Alanine Amidase-(CT268)
CPn0417
           463035
                      462244
                                  R
                                       murE-N-Acetylmuramoylalanylglutamyl DAP Ligase-(CT269)
CPn0418
           464401
                       462953
CPn0419
           466834
                       464876
                                       pbp3- transglycolase/transpeptidase-(CT270)
                                  R
CPn0420
           467108
                       466824
                                  R
                                       CT271 hypothetical protein
CPn0421
           467998
                                       yabC-PBP2B Family methyltransferase-(CT272)
                       467108
                                  R
CPn0422
           458242
                       468784
                                       CT273 hypothetical protein
CPn0423
           468791
                       469216
                                       CT274 hypothetical protein
```

```
dnaA_2-Replication Initiation Factor_2-(CT275)
                                     F
                          470961
   CPn0424
              469612
                                           CT276 hypothetical proteins
                          471564
              470980
   CPn0425
                                           CT277 similarity
                          471536
                                     R
              472111
   CPn0426
                                           nqr2-NADH (Ubiquinone) Dehydrogenase-(CT278)
                                     F
                          473715
              472207
   CPn0427
                                           nqr3-NADH (Ubiquinone) Oxidoreductase, Gamma-(CT279)
                                     F
                          474681
              473722
   CPn0428
                                           nqr4-NADH (Ubiquinone) Reductase 4-(CT280)
                          475319
              474681
   CPn0429
                                           ngr5-NADH (Ubiquinone) Reductase 5-(CT281)
                          476093
   CPn0430
              475326
                                     R
                          476151
              476483
   CPn0431
                          476514
                                     R
   CPn0432
              476816
                                           gcsH-Glycine Cleavage System H Protein-(CT282)
              477273
                          476929
                                     R
   CPn0433
                                           CT283 hypothetical protein
                          477276
   CPn0434
              479462
                                           Phospholipase D superfamily [uncleavable leader peptide]-(CT284)
              480902
                          479475
                                     R
   CPn0435
                                           lplA-Lipoate Protein Ligase-Like Protein-(CT285)
                          480902
                                     R
              481618
   CPn0436
                                           clpC-ClpC Protease-(CT286)
              481816
                          484350
                                     F
  __CPn0437
                                           ycbF-PP-loop superfamily ATPase-(CT287)
                          484334
              485416
   CPn0438
                                     F
              485553
                          486077
   CPn0439
                          486740
                                     F
   CPn0440
              486105
                                           CT007 hypothetical protein
                          487838
                                     F
              486891
   CPn0441
                                           CT006 hypothetical protein
                          488528
                                     F
              488013
   CPn0442
                                           CT005 hypothetical protein
                          489979
                                     F
              488729
   CPn0443
                                           pmp_6-Polymorphic Outer Membrane Protein G/I Family
                          494507
                                     F
              490287
   CPn0444
                                           pmp_7-Polymorphic Outer Membrane Protein G Family
                                     F
              494772
                          497579
   CPn:0445
                                           pmp_8-Polymorphic Outer Membrane Protein G Family
                          500415
                                     F
   CP50446
              497626
                                           pmp_9-Polymorphic Outer Membrane Protein G/I Family
                          503351
                                     F
              500568
   CPn0447
                                           *yxjG_Bs_2 Hypothetical Protein
                          503698
                                     R
               504810
    CP:10448
                                           pmp_10-PMP_10 (Frame-shift with 0451)
                          505330
                                     R
              507231
    CPn0449
                                           pmp_10-Polymorphic Outer Membrane Protein G Family
                          507180
                                     R
               508112
    CPT:0450
                                           pmp_11-Polymorphic Outer Membrane Protein G Family
                                     F
              508275
                          511058
                                           pmp_12-Polymorphic Outer Membrane Protein A/I Family (truncated)
    CPn0451
                                     F
               511319
                          512860
    CPn0452
                                           pmp_13 -Polymorphic Outer Membrane Protein G Family
                                     F
               513234
                          516152
    CPm0453
                                           pmp_14-Polymorphic Outer Membrane Protein H Family
                                     F
               516182
                          519115
    CPn0454
               520348
                          519458
                                      R
    CPn0455
                                      R
               521532
                          520327
    CPn0456
                          522120
                                      R
               523865
    CPT0457
                          524236
                                      R
    CP10458
               526320
               527005
                          526619
    CPn0459
                          526992
    CPn0460
               527840
               528638
                          527844
                                      R
    CPn0461
                          529037
    CPn0462
               531052
                                      R
               532357
                          531191
    CPD0463
                          532366
                                      R
               532842
    CPn0464
               533212
                          532871
                                      R
    CPn0465
                                           pmp_15-Polymorphic Outer Membrane Protein E Family
                          536537
                                      F
               533724
    CPn0466
                                           pmp_16-Polymorphic Outer Membrane Protein E Family
               536633
                          539434
    CPn0467
                                           pmp_17-Polymorphic Outer Membrane Protein E Family
                          540432
                                      F
    CPn0468
               539632
                                           pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0469)
               540399
                          541460
    CPn0469
                                           pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0470)
                          542532
    CPn0470
               541357
                                      P
                                           pmp_18-Polymorphic Outer Membrane Protein E/F Family
    CPn0471
               542564
                          545401
                          545581
                                      R
    CPn0472
               547905
               549593
                           548070
                                      R
    CPn0473
                                           CT365 hypothetical protein
                          549807
    CPn0474
               551573
                                      R
                                           glgB-Glucan Branching Enzyme-(CT866)
               553844
                           551685
                                      R
    CPn0475
                                            CT865 hypothetical protein
                          553858
    CPn0476
               554844
                                      R
                                            *yqeV_Bs Hypothetical Protein
    CPn0477
               556106
                           554844
                                      R
                                           hflx-GTP Binding Protein-(CT379)
    CPn0478
               557625
                           556210
                                      R
                                            phnP-Metal Dependent Hydrolase-(CT380)
    CPn0479
               558425
                           557616
                                      R __CT383_hypothetical_protein_____
-----CPn0480 ---- 559303---- 558650--
               560946
                           559339
                                      R
    CPn0481
                                            artJ-Arginine Periplasmic Binding Protein-(CT381)
    CPn0482
               561737
                           560961
                                      R
                           564964
               561836
    CPn0483
                                            aroG-Deoxyheptonate Aldolase-(CT382)
               564970
                           565824
                                      F
    CPn0484
                                            CT382.1 hypothetical protein
                           566229
                                      F
    CPn0485
               566038
                                            *hypothetical proline permease
    CPn0486
               567784
                           566405
                                      Ŕ
                                            CT384 hypothetical protein
                                      R
    CPn0487
               569740
                           568112
                                            hitA-HIT Family Hydrolase-(CT385)
    CPn0488
               570096
                           569767
                                      R
                                            CT386 hypothetical protein
                                      R
                           570096
    CPn0489
               570965
                                            CT387 hypothetical protein
    CPn0490
               571279
                           573333
                                      F
                                            CT389 hypothetical protein
                                      R
    CPn0491
               574352
                           573336
    CPn0492
               574652
                           574804
                           574855
                                      R
    CPn0493
                575004
    CPn0494
               575364
                           575146
                                      R
                                            aspC-Aspartate Aminotransferase-(CT390)
                           576793
    CPn0495
                575603
```

```
CT391 hypothetical protein
               576793
                          577812
                                     F
   CPn0496
                                          CT388 hypothetical protein
                          577820
                                     R
               578089
   CPn0497
                          578085
                                     R
               579035
   CPn0498
                          579205
               580359
                                     R
   CPn0499
                                          pros-Prolyl tRNA Synthetase-(CT393)
                                     F
               580659
                          582362
   CPn0500
                                          hrcA-HTH Transcriptional Repressor-(CT394)
                                     F
               582457
                          583650
   CPn0501
                                           grpE-HSP-70 Cofactor-(CT395)
                                     F
   CPn0502
               583650
                          584201
                                           dnaK-HSP-70-(CT396)
                                     F
                          586213
               584234
    CPn0503
                                           vacB-ribonuclease family-(CT397)
                                     F
                          588514
               586487
   CPn0504
                                           *3-methyladenine DNA glycosylase
                          589106
                                     F
               588519
    CPn0505
                                           CT421 hypothetical protein
                                     F
                          589840
    CPn0506
               589172
                                           CT421.1 hypothetical protein
                          590122
               589961
    CPn0507
                                           CT421.2 hypothetical protein
                                     F
                          590300
   CPn0508
               590142
                                           (predicted Metalloenzyme) - (CT422)
                          590808
                                     F
               590335
    CPn0509
                                           tlyC_2-CBS Domains (Hemolysin homolog)_2-(CT423)
                                     F
                          591973
               590813
    CPn0510
                                           rsbV_1-Sigma Regulatory Factor_1-(CT424)
                          592488
               592141
                                     F
    CPn0511
                                           CT425 hypothetical protein
                                     F
                          594412
    CPn0512
               592553
                                           Fe-S oxidoreductase_1-(CT426)
                                     F
               594647
                          595753
    CPn0513
                                           CT427 hypothetical protein
                          596520
                                     F
               595729
    CPn0514
                                           ubiE-Ubiquinone Methyltransferase-(CT428)
               596492
                          597181
    CPn0515
                                     R
                          597255
               598814
    CPn0516
                          598795
                                     R
               599631
    CPn0517
                                           CT429 hypothetical protein
    CPn0518
               600803
                          599832
                                     R
                                           dapF-Diaminopimelate Epimerase-(CT430)
                          600904 .
               601674
    CPn0519
                                           clpP-CLP Protease-(CT431)
                                     R
                          601646
    CPn0520
               602218
                                           glyA-Serine Hydroxymethyltransferase-(CT432)
               603797
                          602241
                                     R
    CPn0521
                                           CT433 hypothetical protein
               603987
                          604655
                                     F
    CPn0522
                          605052
    CPn0523
               604723
                          606179
                                     F
               605103
    CPn0524
                                           CT398 hypothetical protein
                          607283
                                     F
    C. 2m.0525
               606522
                                           yrbH-GutQ/KpsF Family Sugar-P Isomerase-(CT399)
                          607710
                                     R
    CPh0526
               608696
                                           sucB_2-Dihydrolipoamide Succinyltransferase_2-(CT400)
               609904
                          608726
    CPn0527
                                           gltT-Glutamate Symport-(CT401)
                          609921
                                     R
    CPn0528
               611162
                                           ycaH-ATPase-(CT402)
                          611165
                                     R
               612259
    CPn0529
                                           spoU_1-rRNA Methylase_1-(CT403)
    CPn0530
               613254
                          612460
                                     R
                                           SAM dependent methyltransferase-(CT404)
                          613245
               614069
    CPn0531
                                           ribC/risA-Riboflavin Synthase-(CT405)
    CPh0532
               614674
                          614075
                                     R
                                           CT406 hypothetical protein
                          615385
                                      F
               614930
    CPn0533
                                           dksA-DnaK Suppressor-(CT407)
                                     F
    CP:n0534
               615413
                          615784
                                           lspA-Lipoprotein Signal Peptidase-(CT408)
                          616296
               615793
    CP50535
                                           dagA_1-D-Ala/Gly Permease_1-(CT409)
                                      F
               616345
                          617691
    CP-0536
                                           CT814.1 hypothetical protein
               617833
                          618189
    CPn0537
                                           CT814 hypothetical protein
                                      F
               618212
                          618511
    CPn0538
                                           pmp_19-polymorphic outer membrane protein A Family -(CT412)
               618705
                          621545
    CPn0539
                                           pmp_20-polymorphic outer membrane protein B Family-(CT413)
                          626862
                                      F
    CPn0540
               621694
                                           Solute binding protein (-yebL-Synechocystis Adhesin Homolog)-(CT415)
               627170
                          628003
                                      F
    CPn0541
                                           ABC Transporter ATPase-(CT416)
                                      F
               628003
                          628737
    CPn0542
                                           (Metal Transport Protein) - (CT417)
    CPn0543
               628725
                          629603
                                      F
                                           yhbZ-GTP binding protein-(CT418)
                                      R
               630529
                          629525
    CPn0544
                                           rl27-L27 ribosomal protein-(CT419)
               630884
                           630633
                                      R
    CPn0545
                                           rl21-L21 Ribosomal Protein-(CT420)
    CPn0546
               631229
                          630912
                                      R
                                           ygbB family-(CT434)
               631661
                          632188
    CPn0547
                                           cysJ-Sulfite Reductase-(CT435)
                          632191
                                      R
    CPn0548
               633231
                                           rs10-S10 Ribosomal Protein-(CT436)
                         • 633255
    CPn0549
               633569
                                      R
                                           fusA-Elongation Factor G-(CT437)
                          633580
    CPn0550
               635661
                                      R
                                           rs7-S7 Ribosomal Protein-(CT438)
               636168
                           635698
                                      R
    CPn0551
                                     R rs12-S12-Ribosomal Protein-(CT439)
---- CPn0552 636587 -- 636219
               637747
                           636812
                                      R
    CPn0553
                                           CT440 hypothetical protein
                637854
                           638141
                                      F
    CPn0554
                                           tsp-Tail-Specific Protease-(CT441)
                638298
                           640241
    CPn0555
                                           crpA-15kDa Cysteine-Rich Protein-(CT442)
                640912
                           640325
                                      R
    CPn0556
                                           omcB-60kDa Cysteine-Rich Outer Membrane Complex Protein-(CT443)
                           641194
                                      R
                642861
    CPn0557
                                           omcA-9kDa-Cysteine-Rich Outer Membrane Complex Lipoprotein-(CT444)
     CPn0558
                643300
                           643031
                                      R
                                            CT441.1 hypothetical protein
                           643927
                643742
    CPn0559
                                            gltX-Glutamyl-tRNA Synthetase-(CT445)
     CPn0560
                645612
                           644098
                                      R
                                            euo-CHLPS Euo Protein-(CT446)
                           645871
                                      R
                646404
     CPn0561
                                            *CHLPS 43 kDa protein homolog_1
     CPn0562
                648036
                           646918
                                      R
                                            recJ-ssDNA Exonuclease-(CT447)
                           648293
                                      R
     CPn0563
                650056
                                            secD&secF-Protein Export Proteins SecD/SecF (fusion)-(CT448)
                           650145
     CPn0564
                654350
                                            CT449 hypothetical protein
                           654533
                                      R
     CPn0565
                655630
                                            yaeS family-(CT450)
                           656890
                                      F
     CPn0566
                656141
                                            cdsA-Phosphatidate Cytidylytransferase-(CT451)
                           657817
     CPn0567
                656894
```

```
cdsA-Phosphatidate Cytidylytransferase-(CT452)
                          658464
              657817
   CPn0568
                                           plsC-Glycerol-3-P Acyltransferase-(CT453)
                          659099
   CPn0569
              658464
                                           argS-Arginyl tRNA Transferase-(CT454)
                          660789
                                     F
              659107
   CPn0570
                                           mura-UDP-N-Acetylglucosamine Transferase-(CT455)
                          660749
   CPn0571
              662122
                                           CT456 hypothetical protein
                          664616
                                     F
              662352
   CPn0572
                                           yebC family-(CT457)
                          664691
                                     R
   CPn0573
              665404
              665945
                          665394
                                     R
   CPn0574
                                           Yhhy-Amino Group Acetyl Transferase-(CT458)
                          665982
              666494
                                     R
                                           prfB-Peptide Chain Release Factor 2 (natural UGA frame-shift )-(CT459
   CPn0575
              667543
                          666494
   CPn0576
                                           prfB-(natural UGA frame-shift )
                          667530
                                     R
              667598
   CPn0576
                                           SWIB (YM74) complex protein-(CT460)
              667895
                          668155
                                     F
   CPn0577
                                           yaeI-phosphohydrolase-(CT461)
                          669365
              668406
   CPn0578
                                           ygbP/yacM-Sugar Nucleotide Phosphorylase-(CT462)
                          669993
              669361
   CPn0579
                                           truA-Pseudouridylate Synthase I-(CT463)
                                     F
              669993
                          670793
   CPn0580
                                           Phosphoglycolate Phosphatase-(CT464)
              671434
                          670745
                                     R
   CPn0581
                                           CT465 hypothetical protein
                          672177
              671503
   CPn0582
                                           CT466 hypothetical protein
                          672717
              672400
   CPn0583
                                           atoS/ntrB-2-Component Sensor-(CT467)
                          673798
                                     F
               672707
   CPn0584
                                           *similarity to Cps IncA_2
                          673865
              675817
   CPn0585
                                           atoC/ntrC-2-Component Regulator-(CT468)
                          677183
                                     F.
               676026
   CPn0586
                                           *yvyD_Bs conserved hypothetical protein
                          678124
                                     F
   CPn0587
               677441
                                           CT469 hypothetical protein
                          678626
                                     F
               678084
   CPn0588
                                           CT470 hypothetical protein
               678640
                          679395
                                     F
   C2a0589
                                           CT471 hypothetical protein
   CPn0590
               680112
                          679516
                                           yagE family-(CT472)
                          681020
               680373
   CPn0591
                                           yidD family-(CT473)
   Ç₽00592
               681153
                          681461
                                     F
                                           CT474 hypothetical protein
               682476
                          681391
    CPn0593
                                           pheT-phenylalanyl tRNA Synthetase Beta-(CT475)
               682583
                          684958
                                     Ŧ
   CPn0594
                                           CT476 hypothetical protein
                          685926
               684958
    CPn0595
                                           ada-methyltransferase-(CT477)
    CPn0596
               685939
                          686457
                                     F
                                           oppC_2-Oligopeptide Permease_2-(CT478)
                          686479
                                      R
               688215
    CPn0597
                                           oppB_2-Oligopeptide Permease_2-(CT479)
                                      R
    CPn0598
               689697
                          688219
                                           oppA_5-oligopeptide Binding Lipoprotein_5-(CT480)
               691802
                          689682
    GPn0599
                                      R
    CPn0600
               692147
                          691827
                                           CT483 hypothetical protein
                          692736
                                      Ŕ
               693053
    CPn0601
                                           CT484 hypothetical protein
               694105
                          693104
                                      R
    ₫₽ħ0602
                                           hemZ-Ferrochetalase-(CT485)
                          695185
               694205
    CPn0603
                                            fliy-Glutamine Binding Protein-(CT486)
                                      R
               695945
                          695196
    CPn0604
                                           yhhF-Methylase -(CT487)
               696707
                          696150
                                      R
    C₽n0605
                                           CT488 hypothetical protein
               697444
                          696707
                                      R
    CPh0606
                                            glgC-Glucose-1-P Adenyltransferase-(CT489)
               698895
                           697573
    CPh0607
                                            *pyrF-Uridine 5 -Monophosphate Synthase (Ump Synthase) -truncated?
               699645.
                           699016
                                      R
    CPn0608
                                            CT490 hypothetical protein
               699705
                           699986
                                      F
    CPn0609
                                            rho-Transcription Termination Factor-(CT491)
               701420
                          700029
                                      R
    CPn0610
                                           yacE-predicted phosphatase/kinase-(CT492)
                           701420
               702025
    CPn0611
                                            polA-DNA Polymerase I-(CT493)
                                      R
               704631
                          702022
    CPn0612
                                            sohB-Protease-(CT494)
                           704658
                                      R
               705656
    CPn0613
                                            adt_2-ADP/ATP Translocase_2-(CT495)
               707402
                           705783
                                      R
    CPn0614
                                            pgsA_1-Glycerol-3-P Phosphatidyltransferase_1-(CT496)
                           707634
    CPn0615
               708137
                                            dnaB-Replicative DNA Helicase-(CT497)
               708791
                           710137
                                      F
    CPn0616
                                            gidA-FAD-dependent oxidoreductase-(CT498)
               710484
                           712316
                                      F
    CPn0617
                                            lplA-Lipoate-Protein Ligase A-(CT499)
    CPn0618
               712306
                           713010
                                      F
                                            ndk-Nucleoside-2-P Kinase-(CT500)
               713444
                           713013
    CPn0619
                                            ruvA-Holliday Junction Helicase-(CT501)
    CPn0620
               714139
                           713519
                                      R
                                            ruvC-Crossover Junction Endonuclease-(CT502)
               714647
                           714144
                                      R
    CPn0621
                                            CT503 hypothetical protein
    CPn0622
               715752
                           714793
                                      R
                                           -CT504-hypothetical-protein-
CPn0623
               -716993
                           7-1:61:63
                                            gapA-Glyceraldehyde-3-P Dehyrogenase-(CT505)
    CPn0624
               718015
                           717011
                                      R
                                            rl17-L17 Ribosomal Protein-(CT506)
               718485
                           718060
                                      R
    CPn0625
                                            rpoA-RNA Polymerase Alpha-(CT507)
    CPn0626
               719616
                           718495
                                      R
                                            rsl1-Sl1 Ribosomal Protein-(CT508)
               720038
                           719640
                                      R
    CPn0627
                                            rs13-S13 Ribosomal Protein-(CT509)
    CPn0628
               720428
                           720063
                                      R
                                            secY-Translocase-(CT510)
                           720487
    CPn0629
                721857
                                            rl15-L15 Ribosomal Protein-(CT511)
                           721885
                                      R
    CPn0630
                22316
                                            rs5-S5 Ribosomal Protein-(CT512)
                           722312
    CPn0631
                722806
                                            rl18-L18 Ribosomal Protein-(CT513)
                           722827
    CPn0632
                723195
                                            rl6-L6 Ribosomal Protein-(CT514)
                           723209
                                       R
    CPn0633
                723757
                                            rs8-S8 Ribosomal Protein-(CT515)
                           723787
                                       R
    CPn0634
                724185
                                            rl5-L5 Ribosomal Protein-(CT516)
                           724206
    CPn0635
                724745
                                       R
                                            rl24-L24 Ribosomal Protein-(CT517)
    CPn0636
                725082
                           724750
                                            rl14-L14 Ribosomal Protein-(CT518)
                           725099
    CPn0637
                725464
                                       R
                                            rs17-S17 Ribosomal Protein-(CT519)
    CPn0638
                725747
                           725490
```

F

```
rl29-L29 Ribosomal Protein-(CT520)
                          725743
              725958
   CPn0639
                                           rl16-L16 Ribosomal Protein-(CT521)
                          725964
   CPn0640
             726377
                                           rs3-S3 Ribosomal Protein-(CT522)
                          726409
              727077
    CPn0641
                                           rl22-L22 Ribosomal Protein-(CT523)
                          727096
   CPn0642
              727428
                                           rs19-S19 Ribosomal Protein-(CT524)
                          727450
              727713
    CPn0643
                                           rl2-L2 Ribosomal Protein-(CT525)
                          727722
              728573
   CPn0644
                                           rl23-L23 Ribosomal Protein-(CT526)
                          728598
              728930
   CPn0645
                                           rl4-L4 Ribosomal Protein-(CT527)
                          728950
                                     R
              729621
    CPn0646
                                           rl3-L3 Ribosomal Protein-(CT528)
                          729657
    CPn0647
              730331
                                           CT529 hypothetical protein
                          730605
                                     R
              731603
    CPn0648
                                           fmt-Methionyl tRNA Formyltransferase-(CT530)
                          731710
                                     R
              732672
    CPn0649
                                           lpxA-Acyl-Carrier UDP-GlcNAc -(CT531)
                          732665
                                     R
               733501
    CPn0650
                                           fabZ-Myristoyl-Acyl Carrier Dehydratase-(CT532)
                          733517
                                     R
               733975
    CPn0651
                                           lpxC-Myristoyl GlcNac Deacetylase-(CT533)
                          733990
                                     R
               734835
    CPn0652
                                           cutE-Apolipoprotein N-Acetyltransferase-(CT534)
                          734868
                                     R
               736490
    CPn0653
                                           vdlD/yciA-acyl-CoA Thioesterase-(CT535)
                          736503
                                     R
               736967
    CPn0654
                                           dnaQ_2-DNA Pol III Epsilon Chain_2-(CT536)
                                     R
                          737101
               737847
    CPn0655
                          738048
                                     F
               737872
    CPn0656
                                           yjeE (ATPase or Kinase)-(CT537)
                          738051
                                     R
               738473
    CPn0657
                                           CT538 hypothetical protein
                          738455
                                     R
               739168
   · CPn0658
                                           trxA-Thioredoxin-(CT539)
                          739838
                                     F
               739533
    CPn0659
                                           spoU_2-rRNA Methylase_2-(CT540)
                          739860
                                      R
               740327
    CPn0660
                                           mip-FKBP-type peptidyl-prolyl cis-trans isomerase-(CT541)
                          740327
                                      R
    CPn0661
               741100
                                           aspS-Aspartyl tRNA Synthetase-(CT542)
               742923
                          741172
                                     R
    CPn0662
                                           hisS-Histidyl tRNA Synthetase-(CT543)
                          742901
                                      R
               744190
    CPn0663
                          744557
               744757
    CPn 0664
                                           uhpC-Hexosphosphate Transport -(CT544)
                                      F
                          746365
    CP:0665
               745001
                                           dnaE-DNA Pol III Alpha-(CT545)
                          750107
                                      F
               746388
    CPm0666
                                           predicted OMP (leader (17)-(CT546)
                          750177
    CPn0667
               751058
                                           CT547 hypothetical protein
    CP#0668
               751209
                          752162
                                           CT548 hypothetical protein
                          752775
                                      F
    CPH0669
               752179
                                           rsbW-sigma regulatory factor-histidine kinase-(CT549)
               752765
                          753196
                                      F
    CPn0670
                                           CT550 hypothetical protein
                          753205
               753630
                                      R
    CPn0671
                                           dacF(pbp5)-D-Ala-D-Ala Caroxypeptidase-(CT551)
               753741
                          755048
    CPn0672
                                           CT552 hypothetical protein
                                      F
                          755463
               755287
    CPn0673
                                           fmu-RNA Methyltransferase-(CT553)
               756668
                           755577
                                      R
    CPT0674
                                           CT696 hypothetical protein
               757919
                           756768
    CP:0675
                                      R
                                           homologous to CT695
    CP:10676
               759217
                           758051
                           759256
    CP00677
               760401
                                      R
    CPn0678
               761320
                           760682
                                           pgk-Phosphoglycerate Kinase-(CT693)
                           761725
               762930
                                      R
    CPn0679
                                           ygo4-Phosphate Permease-(CT692)
               764248
                           762971
    CPn0680
                                           CT691 hypothetical protein
               764929
                           764258
                                      R
    CPn0681
                                            dppD-ABC ATPase Dipeptide Transport-(CT690)
               764984
                           765955
                                      F
    CPn0682
                                           dppF-ABC ATPase Dipeptide Transport-(CT689)
               765948
                           766919
                                      F
    CPn0683
                                            spoJ/parB-Chromosome Partitioning Protein-(CT688)
               768038
                           767181
    CPn0684
    CPn0685
               768068
                           768217
                                      F
               768361
                           768176
                                      R
    CPn0686
                                           CT482 hypothetical protein
    CPn0687
               768564
                           769214
                                      F
                                           CT481 hypothetical protein
               769382
                           770137
    CPn0688
                                           yfhO_1-NifS-related Aminotransferase_1-(CT687)
    CPn0689
               771404
                           770187
                                      R
                                           ABC Transporter Membrane Protein-(CT686)
               772680
                           771436
                                      R
    CPn0690
                                            abcX-ABC Transporter ATPase-(CT685)
    CPn0691
               773452
                           772685
                                      R
                                            ABC Transporter-(CT684)
               774912
                           773461
                                      R
    CPn0692
                                           TPR Repeats (O-Linked GlcNAc Transferase similarity)-(CT683)
    CPn0693
               776256
                           775240
                                      R
                                            pbp2-pBP2-transglycolase/transpeptidase-(CT682)
    CPn0694
               779599
                           776330
                                      R
                                           ompA-Major Outer Membrane-Protein-(CT681)
CPn0695
              780216
                          781382
                                      F
                                            rs2-S2 Ribosomal Protein-(CT680)
               781769
                           782599
    CPn0696
                                            tsf-Elongation Factor TS-(CT679)
    CPn0697
               782602
                           783447
                                      F
                                            pyrH-UMP Kinase-(CT679)
               783458
                           784201
                                      F
    CPn0698
                                            rrf-Ribosome Releasing Factor-(CT677)
    CPn0699
                784182
                           784721
                                      F
                                            CT676 hypothetical protein
                           785609
                                      F
    CPn0700
               785097
                                            karG-Arginine Kinase-(CT675)
    CPn0701
                           786672
                                      F
               785599
                                            yscC/gspD-Yop C/Gen Secretion Protein D-(CT674)
                           786929
                                      R
    CPn0702
               789685
                                            pkn5-S/T Protein Kinase-(CT673)
                           789685
    CPn0703
                791190
                                      R
                                            flin- Flagellar Motor Switch Domain/YscQ family-(CT672)
                           791209
    CPn0704
                792321
                                      R
                                            CT671 hypothetical protein
    CPn0705
                793173
                           792334
                                            CT670 hypothetical protein
    CPn0706
                793683
                           793180
                                      R
                                            yscN-Yop N (Flagellar-Type ATPase) - (CT669)
    CPn0707
                795029
                           793704
                                      R
                                            CT668 hypothetical protein
     CPn0708
                795705
                           795034
                                      R
                                            CT667 hypothetical protein
     CPn0709
                796188
                           795742
                                            CT666 hypothetical protein
     CPn0710
                796461
                           796210
```

```
CT665 hypothetical protein
 CPn0711
            796731
                       796486
                                   R
                                        FHA domain: homology to adenylate cyclase)-(CT664)
            799315
                       796781
                                   R
 CPn0712
                                        CT663 hypothetical protein
                       799332
            799721
                                   R
 CPn0713
                                        hemA-Glutamyl tRNA Reductase-(CT662)
            801107
                       800091
                                   R
 CPn0714
                                        gyrB_2-DNA Gyrase Subunit B_2-(CT661)
                       803462
                                   F
            801657
 CPn0715
                                        gyrA_2-DNA Gyrase Subunit A_2-(CT660)
                       804902
                                   F
            803469
 CPn0716
                                        CT656 hypothetical protein
                       805306
 CPn0717
            805010
                                        CT657 hypothetical protein
                       805626
            805309
 CPn0718
                                        sfhB-(Pseudouridine Synthase)-(CT658)
                       806890
 CPn0719
            805916
                                        CT659 hypothetical protein
            807003
                       807236
 CPn0720
                                        kdsA-KDO Synthetase-(CT655)
                       808489
 CPn0721
            807683
                                        CT654 hypothetical protein
            808489
                       808974
 CPn0722
                                        yhbG-ABC Transporter ATPase-(CT653)
                       809703
                                   F
            808984
 CPn0723
            810527
                       809706
                                   R
CPn0724
                                        CT652.1 hypothetical protein
                       810587
                                   R
            810811
 CPn0725
                                        CT620 hypothetical protein
                       810880
            813372
                                   R
 CPn0726
                                        CT619 hypothetical protein
                       816192
                                   F
 CPn0727
            813577
                                        CHLPN 76kDa Homolog_1 (CT622)
 CPn0728
            818477
                       816525
                                   R
                                        CHLPN 76kDa Homolog_2 (CT623)
                       818592
            819857
                                   R
 CPn0729
                                        mviN-Integral Membrane Protein-(CT624)
 CPn0730
            821603
                       819963
                       821760
                                   F
 CPn0731
            821587
                                        nfo-Endonuclease IV-(CT625)
            822098
                       822976
                                   F
 CPn0732
                                        rs4-S4 Ribosomal Protein-(CT626)
 CP10733
            823727
                       823101
                                   R
                                        yceA-(CT627)
 CPn0734
            823944
                       824915
                                         *pyrH/udk-Uridine Kinase (Uridine Monophosphokinase) (Pyrimidine
 СР 0735
            825668
                       825003
                                   R
                                           Ribonucleoside Kinase).
                                        ygeD-Efflux Protein-(CT641)
                       825992
 CP10736
            827686
                                   R
                                        recC-Exodeoxyribonuclease V, Gamma-(CT640)
            827685
                       830756
                                   F
 CP10737
                                        recB-Exodeoxyribonuclease V, Beta-(CT639)
 CPn0738
            830746
                       833895
                                   F
                                        CT638 hypothetical protein
 CPp0739
            834871
                       833861
                                   R
                                        tyrB-Aromatic AA Aminotransferase-(CT637)
 CP10740
            836048
                       834864
                                   R
                                        greA-Transcription Elongation Factor-(CT636)
 CPH0741
            838350
                       836185
                                        CT635 hypothetical protein
 CPm0742
            838463
                       838888
                                   F
                                        ngrA-Ubiquinone Oxidoreductase, Alpha-(CT634)
 CPm0743
            838962
                       840362
                                   F
                                        hemB-Porphobilinogen Synthase-(CT633)
 CPn0744
            841384
                       840389
                                   R
 CP10745
            841903
                       841742
                                   R
                                        CT632 hypothetical protein
 CPn0746
            841975
                       843567
                                   F
            843675
                       843740
                                   F
                                        CT631 hypothetical protein
 CPn0747
                                        CT631 hypothetical protein (frame-shift)
                                   F
 CPn0747
            843725
                       843910
                                        ispA-Geranyl Transtransferase-(CT628)
 CP110748
            844987
                       844121
                                        glmu-UDP-GlcNAc Pyrophosphorylase-(CT629)
                       845006
 CP10749
            845629
                                   R
                                         tctD/cpxR-HTH Transcriptional Regulatory Protein + Receiver Doman-
 CPn0750
            846411
                       845707
                                   R
                                            (CT630)
                                        CT651 hypothetical protein
            846608
 CPn0751
                       848434
                                   F
                                        recD_2-Exodeoxyribonuclease V, Alpha_2-(CT652)
                       850082
 CPn0752
            848604
                                   F
 CPn0753
            851006
                       850161
                                   R
                                        rs20-S20 Ribosomal Protein-(CT617)
 CPn0754
            851336
                       851040
                                   R
                                        CT616 hypothetical protein
 CPn0755
            851597
                       852799
                                        rpoD-RNA Polymerase Sigma-66 -(CT615)
 CPn0756
            852961
                       854676
                                   F
                                         folX-Dihydroneopterin Aldolase-(CT614)
 CPn0757
            854733
                       855134
                                         folP/dhpS-Dihydropteroate Synthase-(CT613)
 CPn0758
            855110
                       856459
                                   F
                                         folA-Dihydrofolate Reductase-(CT612)
 CPn0759
            856488
                       856997
                                   F
                                        CT611 hypothetical protein
 CPn0760
            856957
                       857694
                                   F
                                        CT610 hypothetical protein
 CPn0761
            857704
                       858375
                                   F
                                        recA-RecA recombination protein-(CT650).
 CPn0762
            859597
                       858539
                                   R
                                        ygfA-Formyltetrahydrofolate Cycloligase-(CT649)
 CPn0763
            860511
                       859972
                                   R
                                        CT648-hypothetical-protein-
CPn0764 - 861807-
                       -860524
                                   R
                                        CT647 hypothetical protein
            862382
                       861801
 CPn0765
                                   R
                                        CT646 hypothetical protein
 CPn0766
            863782
                       862394
                                   R
                                        CT645 hypothetical protein
                       864177
 CPn0767
            863884
                                   F
                                        yohI/nir3-predicted oxidoreductase -(CT644)
 CPn0768
            864159
                       865163
                                   F
                                         topA-DNA Topoisomerase I-Fused to SWI Domain-(CT643)
                       865121
 CPn0769
            867733
                                   R
                                        CT642 hypothetical protein
 CPn0770
            868340
                       869131
                                   F
                                         rpoN-RNA Polymerase Sigma-54-(CT609)
 CPn0771
                       869144
            870463
                                   R
                                         uvrD-DNA Helicase-(CT608)
 CPn0772
            872385
                       870469
                                         ung-Uracil DNA Glycosylase-(CT607)
 CPn0773
                       873195
                                   F
            872488
 CPn0774
            873195
                       873425
                                   F
                                         CT606.1 hypothetical protein
                                        yggV family-(CT606)
 CPn0775
                       873414
            874031
                                   R
                                         CT605 hypothetical protein
 CPn0776
            874246
                       875487
                                         groEL_2-heat shock protein-60 -(CT604)
                       877178
                                   F
 CPn0777
            875601
                                         tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase-(CT603)
 CPn0778
            877505
                       878092
                                   F
                                         CT602 hypothetical protein
                       878095
 CPn0779
            878481
```

```
papQ/amiB-N-Acetylmuramoyl-L-Ala Amidase-(CT601)
               879205
    CPn0780
                                           pal-Peptidoglycan-Associated Lipoprotein-(CT600)
                           879198
                                      R
    CPn0781
               879773
                                            tolB-polysaccharide transporter-(CT599)
                           879773
                                      R
               881065
    CPn0782
                                            CT598 hypothetical protein
                           881100
                                      R
    CPn0783
               881885
                                            exbD-Biopolymer Transport Protein-(CT597)
                           881892
                                      R
               882296
    CPn0784
                                            exbB/tolQ-polysaccharide transporter-(CT596)
                           882296
                                      R
    CPn0785
               882991
                                            dsbD/xprA-Thio:disulfide Interchange Protein-(CT595)
                           885293
                                      F
               883185
    CPn0786
                                           yabD/ycfH-PHP superfamily (urease/pyrimidinase) hydrolase-(CT594)
                                      F
               885619
                           886401
    CPn0787
                                            sdhC-Succinate Dehydrogenase-(CT593)
                           887432
                                      F
               886542
    CPn0788
                                            sdhA-Succinate Dehydrogenase-(CT592)
                                      F
                           889316
    CPn0789
               887439
                                            sdhB-Succinate Dehydrogenase-(CT591)
                           890103
                                      F
               889330
    CPn0790
                                            CT590 hypothetical protein
                                      R
                           890111
    CPn0791
               893050
                                            CT589 hypothetical protein
                           893108
    CPn0792
               894919
                                            rbsU-sigma regulatory family protein-PP2C phosphatase (RsbW
                           894919
                                      R
    CPn0793
               896823
                                              antagonist) - (CT588)
                                      F
               897174
                           898004
    CPn0794
                           899195
               898128
    CPn0795
                                      F
                           901340
    CPn0796
               899301
                           902694
                                      F
    CPn0797
               901600
                           903856
                                      F
               902846
    CPn0798
                           903940
                                      R
    CPn0799
               904986
                                            eno-Enolase-(CT587)
                           905249
                                      R
               906532
    CPn0800
                                            uvrB-Exinuclease ABC Subunit B-(CT586)
                           906727
                                      R
    CPq0801
               908697
                                            trpS-Tryptophanyl tRNA Synthetase-(CT585)
               909740
                           908709
                                      R
    CPn0802
                                            CT584 hypothetical protein
                           909752
                                      R
    CPn0803
               910303
                                            gp6D-CHLTR Plasmid Paralog-(CT583)
                           910310
                                      R
               911059
    CP60804
                                            minD-chromosome partitioning ATPase-CHLTR plasmid protein GP5D-(CT582)
                           911067
                                      R
    CR 0805
               911831
                                            thrs-Threonyl tRNA Synthetase-(CT581)
                           911867
                                      R
    CPn0806
               913771
                                            CT580 hypothetical protein
                           914879
    CPn0807
               913971
                                            CT579 hypothetical protein
               916287
                           914956
                                      R
    CPh0808
                                            CT578 hypothetical protein
                           916307
                                      R
    C#10809
               917785
                                            CT577 hypothetical protein
               918184
                           917825
                                      R
    C-Pn0810
                                            lcrH_1-Low Ca Response Protein H_1-(CT576)
                           918208
                                      R
    CPn0811
                918900
                                            mutL-DNA Mismatch Repair-(CT575)
                919123
                           920862
                                      F
    CPn0812
                                            pepP-Aminopeptidase P-(CT574)
                                      F
                           921934
     CPn0813
                920870
                                            CT573 hypothetical protein
                           923357
                                      F
    CPn0814
                922107
                                            gspD/pilQ-Gen. Secretion Protein D-(CT572)
                           925622
                                      F
    CPn0815
                923361
                                            gspE-Gen. Secretion Protein E-(CT571)
                           927102
                                      F
    CPn0816
                925615
                                            gspF-Gen. Secretion Protein F-(CT570)
                           928287
                                      F
                927115
    CPn0817
                                            predicted OMP [leader (16) peptide]-(CT569)
                           928682
                                      F
    CPn0818
                928314
                                            CT568 hypothetical protein
                           929132
    CPn0819
                928689
                                            CT567 hypothetical protein
                           929659
                                      F
    CPn0820
                929120
                                            CT566 hypothetical protein
                           930668
                                       F
    CPn0821
                929667
                                            CT565 hypothetical protein
                           931229
                                       F
                930756
    CPn0822
                                            yscT/spaR-YopT Tranlocation T-(CT564)
                           931501
                                       R
     CPn0823
                932367
                                            yscS/fliQ-YopS/fliQ Translocation Protein-(CT563)
                           932378
    CPn0824
                932662
                                            yscR-Yop Translocation R-(CT562)
                           932677
                                       R
     CPn0825
                933594
                                            yscL-Yop Translocation L-(CT561)
     CPn0826
                934310
                           933612
                                       R
                                            CT560 hypothetical protein
     CPn0827
                935264
                           934434
                                       R
                                            yscJ-Yop Translocation J-(CT559)
                936271
                           935267
    CPn0828
                                       F
     CPn0829
                936744
                           937298
                937444
                           937959
                                       F
     CPn0830
     CPn0831
                938267
                           938434
                                       F
                                            lipA-Lipoate Synthetase-(CT558)
     CPn0832
                939747
                           938827
                                            lpdA-Lipoamide Dehydrogenase-(CT557)
                           939747
                                       R
     CPn0833
                941129
                                            CT556 hypothetical protein
                           942014
     CPn0834
                941553
                                       F
                                            mot1=1-SWI/SNF-family_helicase_1-(CT555)....
                                      -R----
CPn0835 945689 942045
                                            brnQ-Amino Acid (Branched) Transport-(CT554)
                946879
                           945722
     CPn0836
                                            nth-Enodnuclease III-(CT697)
     CPn0837
                947771
                           947145
                                       R
                                            thdF-Thiophene/Furan Oxidation Protein-(CT698)
                949106
                            947781
                                       R
     CPn0838
                                            psdD-Phosphatidylserine Decarboxylase-(CT699)
     CPn0839
                949257
                           950159
                                       F
                                            CT700 hypothetical protein
                            951544
     CPn0840
                950222
                                            secA_2-Translocase SecA_2-(CT701)
                                       F
     CPn0841
                951731
                            954640
                                            CT702 hypothetical protein (frame-shift with 0843)
     CPn0842
                            954710
                                       R
                954883
                                            CT702 hypothetical protein
                            954994
                                       R
     CPn0843
                955191
                                            yphC-GTPase/GTP-binding protein-(CT703)
                            955270
                                       R
     CPn0844
                956730
                                            pcnB_1-Poly A Polymerase_1-(CT704)
     CPn0845
                958079
                            956850
                                            clpX-CLP Protease ATPase-(CT705)
                            958112
                                       R
     CPn0846
                959374
                                             clpP-CLP Protease Subunit-(CT706)
                                       R
     CPn0847
                959995
                            959387
                                            tig/murI-Trigger Factor-peptidyl-prolyl isomerase-(CT707)
                            960177
     CPn0848
                961502
                                       R
                                            mot1_2-SWI/SNF family helicase_2-(CT708)
                            965285
     CPn0849
                961788
                                            mreB-Rod Shape Protein-Sugar Kinase-(CT709)
     CPn0850
                965293
                            966390
```

878591

R ·

```
CPn0851
             966396
                                         pckA-Phosphoenolpyruvate Carboxykinase-(CT710)
                        968195
  CPn0852
                                         CT711 hypothetical protein
             968316
                        970613
                                   F
                                         CT712 hypothetical protein
  CPn0853
             970637
                         971803
                                   F
                                         ompB-Outer Membrane Protein B-(CT713)
             972837
  CPn0854
                        971806
                                   Я
                                         gpdA-Glycerol-3-P Dehydrogenase-(CT714)
  CPn0855
             973995
                        972994
                                         AgX-1 Homolog-UDP-Glucose Pyrophosphorylase-(CT715)
  CPn0856
             975377
                        973995
             975757
                        975392
                                         CT716 hypothetical protein
  CPn0857
                                   R
                                         fliI-Flagellum-specific ATP Synthase-(CT717)
             977055
                        975757
                                   R
  CPn0858
             977588
                        977055
                                         CT718 hypothetical protein
                                   R
  CPn0859
                                         fliF-Flagellar M-Ring Protein-(CT719)
             978630
                        977608
  CPn0860
             979722
                        978925
                                         nifU-NifU-related protein-(CT720)
  CPn0861
                                   R
             980873
                        979722
                                         yfho_2-NifS-related protein_2-(CT721)
  CPn0862
                                   R
                                         pgmA-Phosphoglycerate Mutase-(CT722)
             981514
                        980831
  CPn0863
                                   R
             981670
                        982374
                                         yjbC-predicted pseudouridine synthase-(CT723)
  CPn0864
                                         CT724 hypothetical protein
             982418
                        982942
                                   F
  CPn0865
             983491
                        982916
                                         birA-Biotin Synthetase-(CT725)
  CPn0866
                                   R
                                         rodA-Rod Shape Protein-(CT726)
             983423
                        984667
                                   F
 CPn0867
                                         zntA/cadA-Metal Transport P-type ATPase-(CT727)
             986643
                        984670
  CPn0868
                                         CT728 hypothetical protein
 CPn0869
             987401
                        986658
                                   F.
                                         serS-Seryl tRNA Synthetase_2-(CT729)
  CPn0870
             988728
                        987448
                                         ribD-Riboflavin Deaminase-(CT730)
             988772
                        989899
 CPn0871
                                   F.
 CPn0872
             989963
                        991216
                                         ribA&ribB-GTP Cyclohydratase & DHBP Synthase -(CT731)
                                         ribE-Ribityllumazine Synthase-(CT732)
             991233
  @Pn0873
                        991694
  CPn0874
             993107
                        991749
                                         CT733 hypothetical protein
 Cen 0875
                                         CT734 hypothetical protein
             993372
                        994022
                                   F
                                         dagA_2-D-Alanine/Glycine Permease_2-(CT735)
 CPn0876
             994144
                        995517
  dp::0877
             995533
                        995982
                                         ybcL family-(CT736)
  CPh0878
             996654
                        995992
                                   F
                                         SET Domain protein-(CT737)
  ÇPîn0879
             997439
                        996645
                                         yycJ-metal dependent hydrolase-(CT738)
                                   R
  CPn0880
             999861
                        997444
                                         ftsK-Cell Division Protein FtsK-(CT739)
                                   R
  ₫₽ħ0881
             1005667
                        1006209
             1006268
 de 10882
                        1007404
                                   F
 CPn0883
             1008865
                        1007573
                                         dmpP/nqr6-Phenolhydrolase/NADH ubiquinone oxidoreductase-(CT740)
                                   R
            1009359
                                         CT741 hypothetical protein
 CPn0884
                        1009009
                                   R
  ÇPn0885
            1010635
                        1009433
                                        ygcA-rRNA Methyltransferse-(CT742)
 CPh0886
            1011276
                                         hctA-Histone-Like Developmental Protein-(CT743)
                        1010908
                                   R
 C2n0887
                                         CHLTR possible phosphoprotein-(CT744)
            1011692
                        1014157
 CPn0888
                                        hemG-protoporphyrinogen Oxidase-(CT745)
            1015423
                        1014119
                                   R
 Cpn0889
                                         hemN_2-Coproporphyrinogen III Oxidase_2-(CT746)
            1016835
                        1015462
 CP-0890
                                        hemE-Uroporphyrinogen Decarboxylase-(CT747)
            1017805
                        1016819
                                   R
 C210891
            1021073
                        1017819
                                   R
                                        mfd-Transcription-Repair Coupling-(CT748)
 CPn0892
            1023661
                        1021046
                                   R
                                        alaS-Alanyl tRNA Synthetase-(CT749)
 CPn0893
            1023894
                        1025888
                                   F
                                        tktB-Transketolase-(CT750)
                                        amn-AMP Nucleosidase-(CT751)
 CPn0894
            1026766
                        1025888
                                   R
 CPn0895
            1026988
                        1027557
                                   F
                                         efp_2-Elongation Factor P_2-(CT752)
 CPn0896
            1027595
                        1027822
                                   F
                                        CT753 hypothetical protein
 CPn0897
            1028737
                        1027853
                                         (possible phosphohydrolase) - (CT754)
                                   R
 CPn0898
            1030460
                        1028904
                                   R
                                        Mitochondrial HSP60 Chaperonin Homolog-(CT755)
 CPn0899
            1030875
                        1032215
                                   F
                                        murF-Muramoyl-DAP Ligase-(CT756)
 CPn0900
            1032235
                        1033281
                                   F
                                        mraY-Muramoyl-Pentapeptide Transferase-(CT757)
 CPn0901
            1033287
                        1034537
                                   F
                                        murD-Muramoylalanine-Glutamate Ligase-(CT758)
 CPn0902
            1034543
                        1035241 .
                                        nlpD-Muramidase (invasin repeat family) - (CT759)
 CPn0903
            1035263
                        1036417
                                        ftsW-Cell Division Protein FtsW-(CT760)
                                   F
 CPn0904
            1036326
                        1037396
                                        murG-Peptidoglycan Transferase-(CT761)
                                   F
 CPn0905
            1037409
                                        murC&ddlA-Muramate-Ala Ligase & D-Ala-D-Alam Ligase-(CT762)
                        1039835
                                   F
 CPn0906
            1040340
                        1039915
                                        CT763 hypothetical protein
           _1040780____1040445__...
                                  R -- CutA Periplasmic Divalent Cation Tolerance Protein CutA (C-Type
CPn0907
                                           Cytochrome Biogenesis Protein)
 CPn0908
            1041589
                        1040780
                                   R
                                        CT764 hypothetical protein
                                        rsbV_2-Sigma Factor Regulator_2-(CT765)
 CPn0909
            1041637
                        1041966
                                   F
 CPn0910
            1041979
                                        miaA-tRNA Pyrophosphate Transferase-(CT766)
                        1043004
                                   F
 CPn0911
            1044043
                        1042985
                                   R
                                        Fe-S cluster oxidoreductase_2-(CT767)
 CPn0912
            1044129
                        1045760
                                   F
                                        CT768 hypothetical protein
 CPn0913
            1045760
                        1045945
                                   F
 CPn0914
            1045999
                        1046397
 CPn0915
            1046461
                                        ybeB-iojap superfamily ortholog-(CT769)
                        1046817
                                   F
 CPn0916
            1046837
                        1048084
                                   F
                                         fabF-Acyl Carrier Protein Synthase-(CT770)
 CPn0917
                                        hydrolase/phosphatase homolog-(CT771)
            1048090
                        1048539
                                   F
 CPn0918
            1049223
                                        ppa-Inorganic Pyrophosphatase-(CT772)
                        1048579
                                   R
 CPn0919
                                        ldh-Leucine Dehydrogenase-(CT773)
            1049378
                        1050430
                                   F
 CPn0920
                                        cysQ-Sulfite Synthesis/biphosphate phosphatase-(CT774)
            1051405
                        1050431
                                   R
 CPn0921
            1051535
                        1052293
                                        snGlycerol-3-P Acyltransferase-(CT775)
```

```
aas-Acylglycerophosphoethanolamine Acydtransferase-(CT776)
 CPn0922
            1052314
                       1053927
                                  F
                                        biof_1-Oxononanoate Synthase_1-(CT777)
 CPn0923
            1053984
                       1055093
                                  F
                                        priA-Primosomal Protein N' -(CT778)
 CPn0924
            1057274
                       1055028
                                        CT779 hypothetical protein
 CPn0925
            1057900
                       1057226
                                  R
            1058060
                       1058557
                                  F
                                        Thioredoxin Disulfide Isomerase-(CT780)
 CPn0926
 CPn0927
            1059809
                       1058670
                                  R
                                        *CHLPS 43 kDa protein homolog_2
            1061008
                       1059884
                                        *CHLPS 43 kDa protein homolog_3
CPn0928
            1062292
                       1061186
                                  R
                                        *CHLPS 43 kDa protein homolog_4
CPn0929
                       1063330
                                  F
            1062857
CPn0930
                                        lyss-Lysyl tRNA Synthetase-(CT781)
CPn0931
            1064138
                       10657.18
                                  F
                                        cysS-Cysteinyl tRNA Synthetase-(CT782)
            1067142
                       1065721
                                  R
CPn0932
CPn0933
            1067535
                       1068578
                                        predicted disulfide bond isomerase-(CT783)
                                        rnpA-Ribonuclease P Protein Component-(CT784)
            1068942
                       1068526
CPn0934
                                  R
CPn0935
            1069091
                       1068957
                                        rl34-L34 Ribosomal Protein-(CT785)
                                        rl36-L36 Ribosomal Protein-(CT786)
            1069336
                       1069470
CPn0936
                                  F
CPn0937
           .1069496
                       1069798
                                        rs14-S14 Ribosomal Protein-(CT787)
                                        CT788 hypothetical protein -[leader (60) peptide-periplasmic]
           1070322
                       1069849
CPn0938
                                  R
                                        CT790 hypothetical protein
CPn0939
            1070728
                       1071195
                                        uvrC-Excinuclease ABC, Subunit C-(CT791)
            1073012
                       1071204
CPn0940
                                  R
                                        mutS-DNA Mismatch Repair-(CT792)
CPn0941
           1075501
                       1073018
                                  R
                                        dnaG/priM-DNA Primase-(CT794)
           1075985
CPn0942
                       1077754
                                  F
CPn0943
           1077978
                       1078238
                                  F
                                        CT794.1 hypothetical protein
CP10944
           1078512
                       1078997
                                  F
CP::0945
           1079070
                       1079660
                                        CT795 hypothetical protein
           1082786
                                        glyQ-Glycyl tRNA Synthetase-(CT796)
CPn0946
                       1079745
                                  R
                                        pgsA_2-Glycerol-3-P-Phosphatydyltransferase_2-(CT797)
CP10947
           1083442
                       1084059
                                  F
CPn0948
           1085474
                       1084047
                                  R
                                        glgA-Glycogen Synthase-(CT798)
CP10949
           1085929
                       1086483
                                        ctc-General Stress Protein-(CT799)
CPn0950
           1086488
                       1087027
                                        pth-Peptidyl tRNA Hydrolase-(CT800)
CPn0951
CPn0952
           1087122
                       1087457
                                        rs6-S6 Ribosomal Protein-(CT801)
                                  F
           1087478
                       1087723
                                        rs18-S18 Ribosomal Protein-(CT802)
           1087742
                                       rl9-L9 Ribosomal Protein-(CT803)
CPn0953
                       1088248
CPn0954
           1088286
                       1088708
                                       ychB-Predicted Kinase-(CT804)
CPn0955
           1088612
                       1089175
                                        (frame-shift with 0954)
CPh0956
           1089560
                       1090909
                                  F
                                        CT805 hypothetical protein
CP10957
                                        ide/ptr-Insulinase family/Protease III-(CT806)
           1093788
                       1090963
                                  R
CPh9958
           1094785
                       1093793
                                       plsB-Glycerol-3-P Acyltransferase-(CT807)
CPh0959
           1096343
                       1094799
                                        cafE-Axial Filament Protein-(CT808)
                                  R
CPn:0960
           1096764
                       1097102
                                  F
                                       CT809 hypothetical protein
CPh0961
           1097118
                       1097297
                                       rl32-L32 Ribosomal Protein-(CT810)
                                  F
CPh 0962
           1097316
                       1098275
                                       plsX-FA/Phospholipid Synthesis Protein-(CT811)
           1098398
                                       pmp_21-Polymorphic Outer Membrane Protein D Family-(CT812)
CPn0963
                       1103224
                                  F
CPn0964
           1104758
                       1103301
                                  R
CPn0965
           1106736
                                       lpxB-Lipid A Disaccharide Synthase-(CT411)
                      1104925
                                  R
CPn0966
           1108037
                       1106748
                                  R
                                       pcnB_2-PolyA Polymerase_2-(CT410)
CPn0967
           1108512
                      1109885
                                  F
                                       mrsA/pgm-Phosphoglucomutase-(CT815)
CPn0968
           1109895
                       1111721
                                  F
                                       glmS-Glucosamine-Fructose-6-P Aminotransferase-(CT816)
CPn0969
           1111812
                      1112999
                                       0969-tyrP_1-Tyrosine Transport_1-(CT817) tyrP_1-Tyrosine Transport_1-
                                  F
                                          (CT817)
CPn0970
           1113461
                      1114648
                                       0970-tyrP_2-Tyrosine Transport_2-(CT818) tyrP_2-Tyrosine Transport_2-
                                  F
                                          (CT818)
CPn0971
           1114702
                      1115415
                                 F
                                       yccA-Transport Permease-(CT819)
CPn0972
           1116299
                      1115430
                                       ftsY-Cell Division Protein FtsY-(CT820)
                                  R
CPn0973
           1116370
                      1117527
                                  F
                                       sucC-Succinyl-CoA Synthetase, Beta-(CT821)
CPn0974
           1117544
                       1118422
                                  F
                                       sucD-Succinyl-CoA Synthetase, Alpha-(CT822)
CPn0975
           1119104
                      1119637
                                  F
CPn0976
           1120082 ____ 1121185
                                  F
CPn0977
           1121371
                      1122402
CPn0978
           1122665
                      1123693
                                  F
CPn0979
           1123980
                      1125443
                                 F
                                       htrA-DO Serine Protease-(CT823)
CPn0980
           1126982
                      1125504
                                       *similarity to Saccharomyces serevisiae hypothetical 52.9KD protein
                                 R
CPn0981
           1127031
                      1129952
                                       Zinc Metalloprotease (insulinase family) - (CT824)
CPn0982
           1131194
                      1129962
                                 R
                                       yigN family-(CT825)
CPn0983
           1132000
                      1131206
                                       pssA-Glycerol-Serine Phosphatidyltransferase-(CT826)
CPn0984
           1132379
                                       nrdA-Ribonucleoside Reductase, Large Chain-(CT827)
                      1135510
CPn0985
           1135534
                                       nrdB-Ribonucleoside Reductase, Small Chain-(CT828)
                      1136571
CPn0986
           1136724
                      1137395
                                       yggH-predicted rRNA Methylase-(CT829)
CPn0987
           1137516
                      1138115
                                 F
                                       ytgB-like predicted rRNA methylase-(CT830)
CPn0988
           1138986
                      1138075
                                 R
                                       murB-UDP-N-Acetylenolpyruvoylglucosamine Reductase-(CT831)
CPn0989
           1139495
                      1139016
                                 R
                                       CT832 hypothetical protein
CPn0990
           1139883
                      1140440
                                       infC-Initiation Factor 3-(CT833)
CPn0991
           1140421
                      1140612
                                       rl35-L35 Ribosomal Protein-(CT834)
```

```
rl20-L20 Ribosomal Protein-(CT835)
                      1140996
                                  F
CPn0992
           1140634
                                        pheS-Phenylalanyl tRNA Synthetase, Alpha-(CT836)
                      1142030
           1141014
                                  F
CPn0993
                                        CT837 hypothetical protein
CPn0994
           1142398
                      1144440
                                  F
                                        CT838 hypothetical protein
                      1144415
                                  R
           1145512
CPn0995
                                        CT839 hypothetical protein
                       1145519
                                  R
CPn0996
           1146589
                                        mesJ-PP-loop superfamily ATPase-(CT840)
                      1147664
CPn0997
           1146708
                                        ftsH-ATP-dependent zinc protease-(CT841)
                       1150584
                                  F
           1147855
CPn0998
                                        pnp-Polyribonucleotide Nucleotidyltransferase-(CT842)
                       1150766
                                  R
           1152847
CPn0999
                                        rs15-S15 Ribosomal Protein-(CT843)
           1153157
                       1152891
                                  R
CPn1000
                                        yfhC-cytosine deaminase-(CT844)
                      1153869
CPn1001
           1153405
                                        CT845 hypothetical protein
                       1154089
                                  F
           1153862
CPn1002
                                        CT846 hypothetical protein
                      1154092
                                  R
CPn1003
           1154796
                                        CT847 hypothetical protein
                       1154879
                                  R
           1155397
CPn1004
                                        CT848 hypothetical protein
                      1155415
                                  R
CPn1005
           1155933
                                        CT849 hypothetical protein
                       1155990
                                  R
           1156472
CPn1006
                                        CT849.1 hypothetical protein
                       1156907
                                  F
CPn1007
           1156689
                                        CT850 hypothetical protein
           1156928
                       1158223
                                  F
CPn1008
                                        map-Methionine Aminopeptidase-(CT851)
                       1158186
                                  R
           1159058
CPn1009
                                        CT852 hypothetical protein
           1159672
                       1159067
CPn1010
                                        CT853 hypothetical protein
                       1159902
                                  R
           1160306
CPn1011
                                        yzeB-ABC transporter permease-(CT854)
           1162193
                       1160421
CPn1012
                                        fumC-Fumarate Hydratase-(CT855)
           1162245
                       1163624
CPn1013
                                        ychM-Sulfate Transporter-(CT856)
           1165426
                       1163732
CPn1-014
                                        CT857 hypothetical protein (possible IM protein)
                      1166893
                                  F
           1165634
CPn1015
                                        CT858 hypothetical protein
           1167042
                       1168898
CPn1016
                                        lytB-Metalloprotease-(CT859)
CPn1017
CPn1018
           1169006
                       1169935
                                  F
                       1170629
           1169898
                                        CT860 hypothetical protein
           1172128
                       1170638
                                  R
CPn 1019
                                        CT861 hypothetical protein
                       1172150
                                  R
CPn1020
           1173679
                                        lcrH_2-Low Calcium Response_2-(CT862)
                       1173698
                                  R
CPn1021
           1174213
                                        CT863 hypothetical protein
                       1174216
CPn 1022
           1175673
           1176035
                       1176331
                                  F
CPn1023
                                        xerD-Integrase/recombinase-(CT864)
                       1176334
                                  R
CPri1024
           1177236
                                        pgi-Glucose-6-P Isomerase-(CT378)
                       1178879
                                  F
CPn 1025
           1177302
                                        ltuA-(CT377)
           1178997
                       1179137
CPn1026
                       1180755
                                  F
CPn1027
           1179175
                                        mdhC-Malate Dehyrogenase-(CT376)
                       1181999
                                   F
           1181016
CPn1028
CPn 1029
           1182008
                       1182844
                                  F
                                        predicted D-amino acid dehyrogenase-(CT375)
                       1182843
                                   R
           1183886
CPn1030
                                        arcD-Arginine/Ornithine Antiporter-(CT374)
CPnF031
           1185552
                       1184098
                                   R
                                        CT373 hypothetical protein
                       1185566
                                   R
CPn1032
           1186150
                                        CT372 hypothetical protein
CPn1033
           1187500
                       1186187
                                   R
                                        Predicted OMP_1 (CT371) [leader (18) peptide]
                       1187732
                                   R
CPn1034
           1188517
                                        AroE-Shikimate 5-Dehyrogenase-(CT370)
                       1188570
CPn1035
           1190000
                                        AroB-Dehyroquinate Synthase-(CT369)
                       1189984
                                   R
CPn1036
           1191135
                                        AroC-Chorismate Synthase-(CT368)
                       1191123
CPn1037
           1192199
                                   R
                                        aroL-Shikimate Kinase II-(CT367)
                       1192199
                                   R
CPn1038
           1192726
                                        aroA-Phosphoshikimate Vinyltransferase-(CT366)
                       1192665
CPn1039
           1193999
                       1194073
                                   R
CPn1040
           1194741
                                         *bioA-Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase
                       1194726
CPn1041
            1195994
                                   R
                                         *bioD-dethiobiotin synthetase
                       1195934
CPn1042
           1196590
                                   R
                                        bioF_2-Oxononanoate Synthase_2
CPn1043
            1197717
                       1196572
                                   R
                                         *bioB-Biotin Synthase
                       1197699
                                   R
CPn1044
            1198691
                                         conserved hypothetical bacterial membrane protein
                       1198901
CPn1045
            1199590
                                         *Tryptophan Hyroxylase
(CPn1046
                       1199590
            1200675
                                         dapB-Dihydrodipicolinate Reductase-(CT364)
 CPn1047
            1200552 -
                       1201343
                                   F
                                         asd-Aspartate_Dehydrogenase-(CT363)
-CPn1048----
           -1201606-
                      --1202604-
                                   F-
                                         lysC-Aspartokinase III-(CT362)
CPn1049
            1202595
                       1203914
                                   F
                                         dapA-Dihydrodipicolinate Synthase-(CT361)
CPn1050
            1203926
                       1204798
                                   F
CPn1051
            1204962
                       1205270
                                   F
- CPn1052
            1205417
                       1206169
                                   F
CPn1053
                       1206701
                                   F
            1206153
CPn1054
            1207034
                       1209466
                                   F
                       1210521
CPn1055
            1209694
CPn1056
            1210527
                        1211228
                                   F
                                   F
                                         CT356 hypothetical protein
CPn1057
                       1213596
            1211497
                                         CT355 hypothetical protein
 CPn1058
            1213748
                        1214836
                                   F
                                         kgsA-Dimethyladenosine Transferase-(CT354)
                       1215678
CPn1059
            1214848
                                   F
                                         dxs/tkt-Transketolase-(CT331)
 CPn1060
            1217658
                       1215727
                                         CT330 hypothetical protein
                                   R
 CPn1061
            1217920
                       1217666
                                         xseA-Exodoxyribonuclease VII-(CT329)
 CPn1062
            1219820
                        1218159
                                         tpiS-Triosephosphate Isomerase-(CT328)
            1219951
                        1220712
 CPn1063
```

Å

CPn1064	1220719	1220895	F	·
CPn1065	1221095	1220928	R	
CPn1066	1221135	1221488	F	
CPn1067	1221735	1222292	F	def-Polypeptide Deformylase-(CT353)
CPn1068	1223258	1222365	R	rnhB_2-Ribonuclease HII_2-(CT008)
CPn1069	1223513	1223941	F	yfgA-HTH Transcriptional Regulator-(CT009)
CPn1070	1225511	1224144	R	
CPn1071	1227324	1225885	R	
CPn1072	1227969	1228835	F	
CPn1073	1229011	1229832	F	Predicted OMP_2 -(CT371)

Table 2 (Supplemental Data) Functional Assignments of *C. pneumoniae* Coding Sequences. *C. trachomatis* genes are shown in parentheses.

	5	Amino Acid Biosynthesis						
		Aromatic	Family					
		1039	(CT366)	aroA	Phosphoshikimate Vinyltransferase			
		1036	(CT369)	aroB	Dehyroquinate Synthase			
		1037	(CT368)	aroC	Chorismate Synthase			
	10	1035	(CT370)	aroÉ	Shikimate 5-Dehyrogenase			
		0484	(CT382)	aroG	Deoxyheptonate Aldolase			
		1038	(CT367)	aroL	Shikimate Kinase II			
		0740	(CT637)	tyrB	Aromatic AA Aminotransferase			
		Aspartate	Family (lys	ine)				
	15	1048	(CT363)	asd	Aspartate Dehydrogenase			
		1050	(CT361)	dapA	Dihydrodipicolinate Synthase			
(I)		1047	(CT364)	dapB	Dihydrodipicolinate Reductase			
١Đ		0519	(CT430)	dapF	Dian inopimelate Epimerase			
2 j es		1049	(CT362)	lysC	Aspa tokinase III			
	20	Serine Fo	imily					
IJ		0433	(CT282)	gcsH	Glyc ne Cleavage System H Protein			
1-1		0521	(CT432)	gìyA	Serine Hydroxymethyltransferase	•		
H)		. Base & N	lucleotide M	etabolism				
M		0171		guaA	GMP Synthase			
1	25	0172		guaB	Inosine 5'-Monophosphase Dehydrogenase			
ļ.		0608			Uridine 5'-Monophosphate Synthase			
ļ.		0735			Uridine Kinase			
ļ		0244	(CT128)	adk	Adenylate Kinase	•		
		0894	(CT751)	amn	AMP Nucleosidase			
E	30	0568	(CT452)	cmk	CMP Kinase			
Han		0392	(CT039)	dcd	dCTP Deaminase			
1-6-7		0059	(CT292)	dut	dUTP Nucleondohydrolase			
		0120	(CT030)	gmk	GMP Kinase			
	25	0619	(CT500)	ndk	Nucleoside-2-P Kinase			
	35	0984	(CT827)	nrdA	Ribonucleoside Reductase, Large Chain			
		0985	(CT828)	nrdB	Ribonucleoside Reductase, Small Chain			
		0236	(CT183)	, pyτG 	CTP Synthetase			
		0698	(CT678)	pyrH 	UMP Kinase			
	40	0273	(CT188)	tdk	Thymidylate Kinase			
	40	0659	(CT539)	trxA	Thioredoxin			
		0314	(CT099)	trxB	Thioredoxin Reductase			
		1001	(CT844)	yfhC	Cytosine Deaminase			
					Biosynthesis of Cofactors			
	45_	Riotin -/	ipoate & Ub	iauinone				
		1041	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase			
		1044		bioB	Biotin Synthase			
		1042		bioD	Dethiobiotin Synthetase			
		0923	(CT777)	bioF_I	Oxononanoate Synthase 1			
	50	1043	(CT777)	bioF_2	Oxononanoate Synthase 2			
			(CT725)	birA	Biotin Synthetase			
		0748	(CT628)	ispA	Geranyl Transtransferase			
		0832	(CT558)	lipA	Lipoate Synthetase			
		0032	(2.330)					

```
0264
                          (CT220)
                                      ubiD
                                                Phenylacrylate Decarboxylase
                   0515
                          (CT428)
                                      ubiE
                                                Ubiquinone Methyltransferase
                 Folic Acid
         5
                                      folA
                                                Dihydrofolate Reductase
                   0759
                          (CT612)
                                                Methylene Tetrahydrofolate Dehydrogenase
                                      folD
                   0335
                          (CT078)
                                                Dihydropteroate Synthase
                          (CT613)
                                      folP
                   0758
                   0757
                          (CT614)
                                      fol X
                                                Dihydroneopterin Aldolase
                                                Formyltetrahydrofolate Cycloligase
                   0763
                          (CT649)
                                     ygfA
       10
                Porphyrin
                                                Glutamyl tRNA Reductase
                   0714
                          (CT662)
                                     hemA
                                                Porphobilinogen Synthase
                   0744
                          (CT633)
                                      hemB
                                                Porphobilinogen Deaminase
                   0052
                          (CT299)
                                      hemC
                          (CT747)
                                      hemE
                                                Uroporphyrinogen Decarboxylase
       15
                   0888
                          (CT745)
                                      hemG
                                                protoporphyrinogen Oxidase
                          (CT210)
                                                Glutamate-1-Semialdehyde-2,1-Aminomutase
                   0138
                                      hemL
                                                Coproporphyrinogen III Oxidase_I
                   0380
                          (CT052)
                                     hemN 1
                   0889
                          (CT746)
                                      hemN_2
                                                Coproporphyrinogen III Oxidase_2
(CT485)
                                                Ferrochetalase
                   0603
                                      hemZ
       20
                 Riboflavin
                   0872 (CT731)
                                      ribA&ribB GTP Cyclohydratase & DHBP Synthase
                   0532
                          (CT405)
                                      пbС
                                                Riboflavin Synthase
                                                Riboflavin Deaminase
                   0871
                          (CT730)
                                      ribD
                   0873
                          (CT732)
                                      пЬЕ
                                                Ribityllumazine Synthase
       25
                   0320
                          (CT093)
                                      пbF
                                                FAD Synthase
M
ñ
                 Cell Envelope
į.
                 Fatty Acid & Phospholipid Metabolism
14
                   0161 (CT206)
                                                (predicted acyltransferase family)
]=
       30
                                                Acylglycerophosphoethanolamine Acyltransferase
                   0922
                         (CT776)
                                     aas
ļ٠
                                                AcCoA Carboxylase/Transferase Alpha
                   0414
                          (CT265)
                                      accA
ū
                                                Biotin Carboxyl Carrier Protein
                   0183
                          (CT123)
                                      accB
Ū
                   0182
                          (CT124)
                                      accC
                                                Biotin Carboxylase
                   0058
                          (CT293)
                                      accD
                                                AcCoA Carboxylase/Transferase Beta
       35
                   0295
                          (CT236)
                                      acpP
                                                Acyl Carrier Protein
                                                Acyl-carrier Protein Synthase
                   0313
                          (CT100)
                                      acpS
                   0567
                          (CT451)
                                      cdsA
                                                Phosphatidate Cytidylytransferase
                   0297
                          (CT238)
                                      fabD
                                                Malonyl Acyl Carrier Transcyclase
                          (CT770)
                                                Acyl Carrier Protein Synthase
                   0916
                                      fabF
       40
                   0296
                          (CT237)
                                      fabG
                                                Oxoacyl (Carrier Protein) Reductase
                                                Oxoacyl Carrier Protein Synthase III
                   0298
                          (CT239)
                                      fabH
                   0406
                          (CT104)
                                      fabl
                                                Enoyl-Acyl-Carrier Protein Reductase
                   0651
                          (CT532)
                                      fabZ
                                                Myristoyl-Acyl Carrier Dehydratase
                   0098
                          (CT010)
                                                Acyltransferase
       45
                   0271
                          (CT136)
                                                Lysophospholipase Esterase
                   0615
                          (CT496)
                                                Glycerol-3-P Phosphatidyltransferase_1
                                      pgsA_1
                   0947
                          (CT797)
                                      pgsA_2
                                                Glycerol-3-P Phosphatydyltransferase_2
                                                Glycerol=3=P=Acyltransferase=
                          -(CT807)-
                                     :plsB=
                                                Glycerol-3-P Acyltransferase
                   0569
                          (CT453)
                                     pisC
       50
                                                FA/Phospholipid Synthesis Protein
                   0962
                          (CT811)
                                      plsX
                   0839
                          (CT699)
                                      psdD
                                                Phosphatidylserine Decarboxylase
                                                Glycerol-Serine Phosphatidyltransferase
                   0983
                          (CT826)
                                      pssA
                                                snGlycerol-3-P Acyltransferase
                   0921
                          (CT775)
                   0654
                                                Acyl-CoA Thioesterase
                          (CT535)
                                      yciA
       55
                   0877
                         (CT736)
                                                CT736 Hypothetical Protein
                                      ybcL
                 LPS
```

(CT219)

ubiA

0265

Benzoate Octaphenvitransferase

```
0154 (CT208)
                                            KDO Synthetase
                                  kdsA
                0721
                       (CT655)
                                            Deoxyoctulonosic Acid Synthetase
                 0235
                       (CT182)
                                  kdsB
                                            Acyl-Carrier UDP-GlcNAc O-Acyltransferase
                                  lpxA
                 0650
                       (CT531)
       5
                                  lpxB
                                            Lipid A Disaccharide Synthase
                       (CT411)
                 0965
                                            Myristoyl GlcNac Deacetylase
                                  lpxC
                       (CT533)
                 0652
                                            UDP Glucosamine N-Acyltransferase
                                  ipxD
                 0302
                       (CT243)
              Membrane Proteins, Lipoproteins & Porins
                                            60kDa Inner Membrane Protein
                 0310
                       (CT251)
                                  60IM
      10
                                  стрА
                                            15kDa Cysteine-Rich Protein
                 0556
                       (CT442)
                                            Apolipoprotein N-Acetyltransferase
                                  cutE
                 0653
                       (CT534)
                                            Prolipoprotein Diacylglycerol Transferase
                 0311
                       (CT252)
                                  lgt
                                            9kDa-Cysteine-Rich Lipoprotein
                 0558
                       (CT444)
                                  omcA
                                            60kDa Cysteine-Rich OMP
                       (CT443)
                                  omcB
                                            Major Outer Membrane Protein
      15
                       (CT681)
                                  ompA
                 0695
                                            Outer Membrane Protein B
                       (CT713)
                                  ompB
                 0854
                                            Peptidoglycan-Associated Lipoprotein
                       (CT600)
                                  pal
                 0781
                                            Omp85 Homolog
                 0300
                       (CT241)
                                  yaeT
               Peptidoglycan
N-Acetylmuramoyl Alanine Amidase
      20
                 0417 (CT268)
                                  amiA
                                            N-Acetylmuramoyl-L-Ala Amidase
                 0780
                       (CT601)
                                  amiB
                                            D-Ala-D-Ala Caroxypeptidase
                 0672
                        (CT551)
                                  dacF
                                            Glucosamine-Fructose-6-P Aminotransferase
                        (CT816)
                                   glmS
                 0968
Ö
                                            UDP-GlcNAc Pyrophosphorylase
                  0749
                        (CT629)
                                   glmU
                                            Muramoyi-Pentapeptide Transferase
      25
(CT757)
                                   mraY
                                            UDP-N-Acetylglucosamine Transferase
IJ
                  0571
                        (CT455)
                                   murA
                                             UDP-N-Acetylenolpyruvoylglucosamine Reductase
                        (CT831)
                                   murB
M
                  0988
                                   murC&ddlA Muramate-Ala Ligase & D-Ala-D-Alam Ligase
                  0905
                        (CT762)
                                             Muramoylalanine-Glutamate Ligase
                  0901
                        (CT758)
                                   murD
1
                                             N-Acetylmuramoylalanylglutamyl DAP Ligase
       30
                        (CT269)
                                   murE
                                             Muramoyl-DAP Ligase
                  0899
                        (CT756)
                                   murF
Peptidoglycan Transferase
                        (CT761)
                                   murG
                  0904
Muramidase (invasin repeat family)
                        (CT759)
                                   nlpD
                  0902
ű
                                             PBP2-Transglycolase/Transpeptidase
                  0694
                        (CT682)
                                   pbp2
ū
       35
                                             Transglycolase/Transpeptidase
                  0419
                        (CT270)
                                   pbp3
                        (CT272)
                                   yabC
                                             PBP2B Family Methyltransferase
                  0421
                                                                        Cellular Processes
                Cell Division
       40
                                             Axial Filament Protein
                  0959
                       (CT808)
                                   cafE
                                   ftsK
                                             Cell Division Protein FtsK
                  0880
                        (CT739)
                        (CT760)
                                   ftsW
                                             Cell Division Protein FtsW
                  0903
                                             Cell Division Protein FtsY
                  0972
                        (CT820)
                                    ftsY
                                             FAD-dependent Oxidoreductase
                  0617
                        (CT498)
                                    gidA
       45
                                             Chromosome Partitioning ATPase
                  0805
                        (CT582)
                                    minD
                  0850
                        (CT709)
                                    mreB
                                             Rod Shape Protein-Sugar Kinase
                  0867
                        (CT726)
                                    rodA
                                             Rod Shape Protein
                  -0684 =(CT688) --- parB-
                                             Chromosome: Partitioning: Protein
                Detoxtification
       50
                                             Superoxide Dismutase (Mn)
                   0057 (CT294)
                                    sodM
                                             Thio-specific Antioxidant (TSA) Peroxidase
                                    ahpC
                  0778 (CT603)
                Signal Transduction
                                              S/T Protein Kinase
                  0148 (CT145)
                                             Two-Component Sensor
                   0584
                         (CT467)
                                    atoS
       55
                                              cAMP-Dependent Protein Kinase Regulatory Subunit
                   0294
                         (CT235)
                   0712 (CT664)
                                              (FHA domain)
```

KDO Transferase

gseA

```
GTP Binding Protein
                   0478 (CT379)
                                    hflX
                                               S/T Protein Kinase
                   0703
                         (CT673)
                   0095
                         (CT301)
                                               S/T Protein Kinase
                   0397
                          (CT259)
                                               PP2C Phosphatase Family
         5
                                               PTS Phosphocarrier Protein Hpr
                   0037
                         (CT337)
                                     ptsH
                                               PTS PEP Phosphotransferase
                   0038
                          (CT336)
                                     ptsI
                                               PTS IIA Protein_I
                   0060
                          (CT291)
                                     ptsN_1
                                               PTS IIA Protein + HTH DNA-Binding Domain
                                     ptsN_2
                   0061
                          (CT290)
                                     surE
                                                SurE-like Acid Phosphatase
                   0262
                         (CT218)
                                                Thiophene/Furan Oxidation Protein
       10
                                     thdF
                   0838
                         (CT698)
                                                TPR Repeats-CT683 Hypothetical Protein
                   0693
                          (CT683)
                                                GTP Binding Protein
                          (CT092)
                                     ychF
                   0321
                                     yhbZ
                                                GTP binding protein
                   0544
                          (CT418)
                                                GTPase/GTP-binding protein
                          (CT703)
                                     yphC
                   0844
       15
                Standard Protein Secretion
                                                Signal Recognition Particle GTPase
                          (CT025)
                   0115
                                     ffh
                                                Flagellar Secretion Protein
                                     flhA
                   0363
                          (CT060)
                                               Flagellum-specific ATP Synthase
                   0858
                          (CT717)
                                      fliI
                                               Flagellar Motor Switch Domain/YscQ family
                   0704
                          (CT672)
                                      fliN
20
                                                Gen. Secretion Protein D
                   0815
                          (CT572)
                                      gspD
                                                Gen. Secretion Protein E
                   0816
                          (CT571)
                                      gspE
                                                Gen. Secretion Protein F
                   0817
                          (CT570)
                                      gspF
                          (CT064)
                                                GTPase
                   0359
                                      lepA
                                                Signal Peptidase I
                   0110
                          (CT020)
                                      lepB
14
       25
                                                Lipoprotein Signal Peptidase
                   0535
                          (CT408)
                                      IspA
                                                Protein Translocase Subunit_1
Ш
                   0260
                          (CT141)
                                      secA 1
                                                Translocase SecA 2
M
                    0841
                          (CT701)
                                      secA_2
                                                  Protein Export Proteins SecD/SecF (fusion)
                                      secD&secF
                    0564
                          (CT448)
                                                Preprotein Translocase
                    0075
                          (CT321)
                                      secE
       30
                                                Translocase
                    0629
                          (CT510)
                                      sec Y
 ŀ
                                                Trigger Factor-Peptidyl-prolyl Isomerase
                    0848
                          (CT707)
                                      tig
 ļ=
                 Transport-Related Proteins
 H
                                                Hypothetical Proline Permease
                    0486
 ١Ū
                                                Neutral Amino Acid (Glutamate) Transporter
                    0289
                          (CT230)
                                      aaaT
 Q
        35
                                                ABC Transporter ATPase
                    0691
                          (CT685)
                                      abcX
                                                Arginine/Omithine Antiporter
                                      arcD
                    1031
                          (CT374)
                                                Arginine Periplasmic Binding Protein
                                      arti
                    0482
                          (CT381)
                                                Amino Acid (Branched) Transport
                    0836
                          (CT554)
                                      bmQ
                                                D-Ala/Gly Permease_l
                    0536
                           (CT409)
                                      dagA_1
        40
                                                D-Alanine/Glycine Permease_2
                    0876
                          (CT735)
                                      dagA_2
                                                ABC ATPase Dipeptide Transport
                    0682
                                      dppD
                          (CT690)
                                                ABC ATPase Dipeptide Transport
                    0683
                          (CT689)
                                      dppF
                                                Dipeptide Transporter ATPase
                    0280
                           (CT689)
                                      dppF
                           (CT596)
                                      exbB
                                                Macromolecule Transporter
                    0785
        45
                                                 Biopolymer Transport Protein
                    0784
                           (CT597)
                                      exbD
                                      fliY
                                                 Glutamine Binding Protein
                    0604
                           (CT486)
                                                 ABC Amino Acid Transporter Permease
                    0192
                          (CT129)
                                      glnP
                                      glnQ
                                                 ABC Amino Acid Transporter ATPase
                           (CT130)_
                    0191_
                                                 Glutamate Symport
                    0528
                           (CT401)
                                      gltT
        50
                                      mgtE
                                                 Mg++ Transporter (CBS Domain)
                    0286
                           (CT194)
                                                 Transport ATP Binding Protein
                    0413
                           (CT264)
                                      msbA
                                                 Na+-dependent Transporter
                    0290
                           (CT231)
                                                 Oligopeptide Binding Protein_1
                    0195
                           (CT198)
                                      oppA 1
                                                 Oligopeptide Binding Protein_2
                    0196
                           (CT198)
                                      oppA_2
        55
                                                 Oligopeptide Binding Protein_3
                    0197
                           (CT139)
                     0198
                           (CT175)
                                       oppA_4
                                                 Oligopeptide Binding Protein 4
```

```
Oligopeptide Permease_1
                   0200
                         (CT200)
                                     oppC_1
         5
                                               Oligopeptide Permease_2
                   0597
                         (CT478)
                                     oppC_2
                                               Oligopeptide Transport ATPase
                   0201
                         (CT201)
                                     OppD
                                               Oligopeptide Transport ATPase
                         (CT202)
                   0202
                                     oppF
                   0231
                         (CT180)
                                     tauB
                                               ABC Transport ATPase (Nitrate/Fe)
                   0782
                         (CT599)
                                     tolB
                                              Macromolecule Transporter
       10
                                               Tyrosine Transport_1
                   0969
                         (CT817)
                                     tyrP_1
                                              Tyrosine Transport 2
                   0970
                         (CT8,18)
                                     tyrP 2
                                               Hexosphosphate Transport
                   0665
                         (CT544)
                                     uhpC
                   0282
                         (CT216)
                                               Amina Acid Transporter
                                     xasA
                   0207
                                     ybhl
                                              dicarboxylate Translocator
                         (CT204)
       15
                                              Transport Permease
                   0971
                         (CT819)
                                     yc¢A
                   0248
                         (CT152)
                                     ycfV
                                               ABC Transporter ATPase
                   1014
                         (CT856)
                                     ychM
                                              Sulfa e Transporter
                                               Efflu: Protein
                   0736
                         (CT641)
                                     ygeD
Phospitate Permease
                   0680
                         (CT692)
                                     ygo4
       20
                   0723
                         (CT653)
                                              ABC Transporter ATPase
                                     yhbG
                   0023
                         (CT348)
                                     уijК
                                              ABC Transporter Protein ATPase
                                              Cationic Amino Acid Transporter
                   0127
                         (CT034)
                                     ytfF
                                              Solut: Protein Binding Family
                   0349
                         (CT067)
                                     ytgA
ŀ÷
                                               ABC transporter ATPase
                   0348
                         (CT068)
                                     ytgB
       25
U
                   0347
                         (CT069)
                                     ytgC
                                               Integral Membrane Protein
                                              Integral Membrane Protein
                   0346
                         (CT070)
                                     ytgD
LM
                                              ABC Transporter Permease
                   1012
                         (CT854)
                                     yzeB
                   0868
                         (CT727)
                                              Metal Transport P-type ATPase
                                     zntA
Ŀ
                   0279
                                              Possible ABC Transporter Permease Protein
14
       30
                   0543
                         (CT417)
                                              (Metal Transport Protein)
14
                   0692
                         (CT684)
                                              ABC Transporter
ļ÷
                                              ABC Transporter ATPase
                   0542
                         (CT416)
ŧŪ
                         (CT686)
                                              ABC Transporter Membrane Protein
                   0690
١Ď
                   0541
                         (CT415)
                                              solute binding protein
       35
                Type-III Secretion
                   0323 .(CT090)
                                     lcrD
                                              Low Calcium Response D
                                              Low Calcium Response E
                   0324
                         (CT089)
                                     IcrE
                                              Low Ca Response Protein H 1
                   0811
                         (CT576)
                                     lcrH_1
                                     lcrH_2
                   1021
                                              Low Calcium Response_2
                         (CT862)
       40
                   0325
                         (CT088)
                                     sycE
                                              Secretion Chaperone
                   0702
                                              Yop C/Gen Secretion Protein D
                         (CT674)
                                     yscC
                   0828
                         (CT559)
                                              You Translocation J
                                     vscJ
                   0826
                         (CT561)
                                     yscL
                                              Yop Translocation L
                                    yscN
                   0707
                         (CT669)
                                              Yop N (Flagellar-Type ATPase)
       45
                   0825
                         (CT562)
                                    yscR
                                              Yop Translocation R
                   0824
                         (CT563)
                                    yscS
                                              YopS Translocation Protein
                   0823
                         (CT564)
                                              YopT Tranlocation T
                                     yscT
                   0322
                         (CT091)
                                     yscU
                                               Yop Translocation Protein U
       50
                                                                   Central Intermediary Metabolism
                Glycogen Metabolism
                   0856 (CT715)
                                              UDP-Glucose Pyrophosphorylase
                   0948
                         (CT798)
                                              Glycogen Synthase
                                     glgA
                   0475
                         (CT866)
                                     glgB
                                              Glucan Branching Enzyme
       55
                   0607
                         (CT489)
                                     glgC
                                              Glucose-I-P Adenyltransferase
                   0307
                         (CT248)
                                     glgP
                                              Glycogen Phosphorylase
                         (CT042)
                                              Glycogen Hydrolase (debranching)
                   0388
                                     glgX
```

Oligopeptide Binding Lipoprotein_5

Oligopeptide Permease_1

Oligopeptide Permease_2

0599

0199

0598

(CT480)

(CT199)

(CT479)

oppA_5

oppB_I

oppB_2

```
0326 (CT087)
                                             Glucanotransferase
                                   malO
                 0851 (CT710)
                                   pckA
                                             Phosphoenolpyruvate Carboxykinase
               Phosphorous & Sulfur
                 0548 (CT435)
                                   cysJ
                                             Sulfite Reductase
        5
                       (CT774)
                                             Sulfite Synthesis/Biphosphate Phosphatase
                 0920
                                   cysQ
                       (CT346)
                                             Sulphohydrolase
                 0025
                                   atsA
                                             Inorganic Pyrophosphatase
                 0918 (CT772)
                                   ppa
                                                       DNA Replication, Modification, Repair & Recombination
      10
               DNA Mismatch Repair
                                             3-Methyladenine DNA Glycosylase
                 0505
                                             DNA Mismatch Repair
                 0812
                        (CT575)
                                   mutL
                 0941
                        (CT792)
                                   mutS
                                             DNA Mismatch Repair
                                             Adenine Glycosylase
                 0402
                        (CT107)
                                   mutY
      15
                                             Endonuclease IV
                        (CT625)
                                   nfo
                 0732
                                             Enodnuclease III
                 0837
                        (CT697)
                                   nth
               DNA Modification
                 0596
                        (CT477)
                                   ada
                                             Methyltransferase
                        (CT024)
                                             A/G-specific Methylase
                 0114
                                   hemK
20
                        (CT748)
                                   mfd
                                             Transcription-Repair Coupling
                 0891
                                             Holliday Junction Helicase
                 0620
                        (CT501)
                                   ruvA
                                             Holliday Junction Helicase
                 0390
                        (CT040)
                                   πινΒ
                        (CT502)
                                   ruvC
                                             Crossover Junction Endonuclease
                  0621
                                             Sms Protein
                  0053
                        (CT298)
                                   sms
ļ
      25
                                             Uracil DNA Glycosylase
                  0773
                        (CT607)
                                   ung
Ø
                                             Exodoxyribonuclease VII
                  1062
                        (CT329)
                                   xseA
n
               DNA Recombination
Ē
                  0762 (CT650)
                                   recA
                                             RecA Recombination Protein
1=1
                  0738
                        (CT639)
                                   recB
                                             Exodeoxyribonuclease V, Beta
14
      30
                                             Exodeoxyribonuclease V, Gamma
                  0737
                        (CT640)
                                   recC
14
                                   recD_1
                                             Exodeoxyribonuclease V (Alpha Subunit)_1
                 0123
                        (CT033)
14
                                             Exodeoxyribonuclease V, Alpha_2
                                   recD_2
                  0752
                        (CT652)
Q
                  0339
                        (CT074)
                                   recF
                                              ABC Superfamily ATPase
·
                  0340
                        (CT074)
                                             (frame-shift with 0339)
      35
                        (CT447)
                  0563
                                             ssDNA Exonuclease
                                    recJ
                        (CT240)
                                             Recombination Protein
                  0299
                                    recR
               DNA Replication
                  0309
                        (CT250)
                                             Replication Initiation Protein_1
                                   dnaA_1
                  0424
                        (CT275)
                                   dnaA_2
                                             Replication Initiation Factor_2
      40
                  0616
                        (CT497)
                                   dnaB
                                              Replicative DNA Helicase
                        (CT545)
                                   dnaE
                                              DNA Pol III Alpha
                  0666
                        (CT794)
                                              DNA Primase
                  0942
                                    dnaG
                                              DNA Pol III (Beta)
                  0338
                        (CT075)
                                    dnaN
                  0410
                        (CT261)
                                    dnaQ_i
                                              DNA Pol III Epsilon Chain_1
      45
                        (CT536)
                                              DNA Pol III Epsilon Chain_2
                  0655
                                   dnaQ_2
                        (CT334)
                                              DNA Pol III Gamma and Tau_I
                  0040
                                    dnaX_i
                  0272
                        (CT187)
                                    dnaX_2
                                              DNA Pol III Gamma and Tau_2
                                              DNA Ligase
                 =0149
                        (CT146)=
                                  -- dull -- -
                  0274
                        (CT189)
                                              DNA Gyrase Subunit A_1
                                    gyrA_1
      50
                  0716
                        (CT660)
                                              DNA Gyrase Subunit A_2
                                    gyrA_2
                                              DNA Gyrase Subunit B_I
                  0275
                        (CT190)
                                    gyrB_1
                  0715
                        (CT661)
                                    gyrB_2
                                              DNA Gyrase Subunit B_2
                                              Integration Host Factor Alpha
                  0416
                        (CT267)
                                    himD
                  0612
                        (CT493)
                                    polA
                                              DNA Polymerase I
      55
                  0924
                        (CT778)
                                    priA
                                              Primosomal Protein N'
                  0386
                        (CT044)
                                              SS DNA Binding Protein
```

```
SWI/SNF family helicase 2
                  0849
                        (CT708)
                                             DNA Topoisomerase I-Fused to SWI Domain
                  0769
                        (CT643)
                                   topA
                  0024
                        (CT347)
                                   xerC
                                             Integrase/recombinase
        5
                                             Integrase/recombinase
                  1024
                        (CT864)
                                   xerD
               Eukaryotic-Type Chromatin Factors
                                             Histone-Like Developmental Protein
                  0886 (CT743)
                                   hctA
                                             Histone-like Protein 2
                                   hctB
                  0384 (CT046)
                        (CT737)
                                             SET Domain protein
      10
                                             SWIB (YM74) Complex Protein
                  0577
                        (CT460)
               UVR Exinuclease Repair System
                                             Excinuclease ABC Subunit A
                        (CT333)
                  0096
                                   uvrA
                                             Exinuclease ABC Subunit B
                                    uvrB
                  0801
                        (CT586)
                                             Excinuclease ABC, Subunit C
                  0940
                        (CT791)
                                   uvrC
      15
                  0772
                        (CT608)
                                    uvrD
                                              DNA Helicase
                                                                        Energy Metabolism
               Aerobic
                  0855
                        (CT714)
                                    gpdA
                                              Glycerol-3-P Dehydrogenase
      20
0743
                         (CT634)
                                              Ubiquinone Oxidoreductase, Alpha
                                    ngrA
                                   nqr2
                                              NADH (Ubiquinone) Dehydrogenase
                  0427
                         (CT278)
                                              NADH (Ubiquinone) Oxidoreductase, Gamma
                  0428
                         (CT279)
                                    nqr3
                  0429
                         (CT280)
                                    nqr4
                                              NADH (Ubiquinone) Reductase 4
Ø
                                              NADH (Ubiquinone) Reductase 5
                  0430
                        (CT281)
                                    nqr5
      25
                                              Phenolhydrolase/NADH (Ubiquinone) Oxidoreductase 6
ļ.
                  0883
                        (CT740)
                                    naró
П
               ATP Biogenesis and metabolism
                                              ADP/ATP Translocase_1
I
                  0351
                         (CT065)
                                    adt_l
                         (CT495)
                                    adt_2
                                              ADP/ATP Translocase_2
#
                                              ATP Synthase Subunit A
                  0088
                         (CT308)
                                    atpA
1-4
      30
                                   atpB
                                              ATP Synthase Subunit B
                  0089
                         (CT307)
14
                                              ATP Synthase Subunit D
                         (CT306)
                  0090
                                    atpD
ļ.
                  0086
                        (CT310)
                                    aφE
                                              ATP Synthase Subunit E
i di
                         (CT305)
                                              ATP Synthase Subunit I
                  0091
                                    atpl
·D
                  0092
                        (CT304)
                                    atpK
                                              ATP Synthase Subunit K
·D
      35
                  0860 (CT719)
                                    fliF
                                              Flagellar M-Ring Protein
               Electron Transport Chain
                                              Cytochrome Oxidase Subunit I
                  0102 (CT013)
                                    cydA
                                              Cytochrome Oxidase Subunit II
                  0103 (CT014)
                                    cydB
                  0364 (CT059)
                                              Ferredoxin
      40
                                              Predicted Ferredoxin
                  0084 (CT312)
                Glycolysis & Gluconeogenesis
                  0281
                         (CT215)
                                              Predicted 1,6-Fructose Biphosphate Aldolase
                                    dhnA
                                              Enolase
                  0800
                         (CT587)
                                    eno
                  0624
                         (CT505)
                                    gapA
                                              Glyceraldehyde-3-P Dehyrogenase
       45
                  0056
                         (CT295)
                                              Phosphomannomutase
                                    mrsA
                  0967
                        (CT815)
                                              Phosphoglucomutase
                                    pgm
                                              Fructose-6-P Phosphotransferase_1
                  0160
                        (CT207)
                                    pfkA_1
                  0208
                        ~(CT205)
                                    pfkA-2
                                              Fructose=6-P Phosphotransferase_2
                   1025
                         (CT378)
                                              Glucose-6-P Isomerase
                                    pgi
       50
                         (CT693)
                                              Phosphoglycerate Kinase
                  0679
                                    pgk
                                              Phosphoglycerate Mutase
                  0863 (CT722)
                                    pgmA
                  0097
                        (CT332)
                                    pyk
                                              Pyruvate Kinase
                                              Triosephosphate Isomerase
                  1063
                        (CT328)
                                    tpiS
                Pentose Phosphate Pathway
       55
                                              Glucose-6-P Dehyrogenase (DevB family)
                  0239 (CT186)
                                    devB
                                              Transketolase
                   1060 (CT331)
                                    dxs
```

SWI/SNF family helicase_1

0835

(CT555)

```
Ribulose-P Epimerase
                  0185
                         (CT121)
                                    тре
                                              Ribose-5-P Isomerase A
                         (CT213)
                                    тріA
                  0141
                                              Transaldolase
                  0083
                         (CT313)
                                    tal
        5
                  0893
                         (CT750)
                                    ιktΒ
                                              Transketolase
                                              Glucose-6-P Dehyrogenase
                  0238
                        (CT185)
                                    zwf
               Pyruvate Dehydrogenase
                        (CT557)
                                    IpdA
                                              Lipoamide Dehydrogenase
                  0833
                                              Lipoate Protein Ligase-Like Protein
                        (CT285)
                                    lplA_l
                  0436
      10
                  0618
                         (CT499)
                                    lplA_2
                                              Lipoate-Protein Ligase A
                        (CT340)
                                    pdhA&B
                                              Oxoisovalerate Dehydrogenase α/β Fusion
                  0033
                                              Pyruvate Dehydrogenase Alpha
                        (CT245)
                                    pdhA
                  0304
                                              Pyruvate Dehydrogenase Beta
                  0305
                         (CT246)
                                    pdhB
                        (CT247)
                                    pdhC
                                              Dihydrolipoamide Acetyltransferase
                  0306
      15
               TCA Cycle
                  0495 (CT390)
                                    aspC
                                              Aspartate Aminotransferase
                  1013
                         (CT855)
                                    fumC
                                              Furnarate Hydratase
                  1028
                         (CT376)
                                    mdhC
                                              Malate Dehyrogenase
                  0789
                         (CT592)
                                    sdhA
                                              Succinate Dehydrogenase
20
                  0790
                         (CT591)
                                    sdhB
                                              Succinate Dehydrogenase
                                              Succinate Dehydrogenase
                                    sdhC
                  0788
                         (CT593)
                                              Oxoglutarate Dehydrogenase
                  0378
                         (CT054)
                                    sucA
                  0377
                         (CT055)
                                    sucB_1
                                              Dihydrolipoamide Succinyltransferase_1
                                              Dihydrolipoamide Succinyltransferase_2
                  0527
                         (CT400)
                                    sucB_2
      25
                  0973
                         (CT821)
                                    sucC
                                              Succinyl-CoA Synthetase, Beta
                                              Succinyl-CoA Synthetase, Alpha
                  0974
                        (CT822)
                                    sucD
U
Ħ.
                                                               Protein Folding, Assembly & Modification
ļ±
                Chaperones
ļ.
       30
                  0949 (CT799)
                                              General Stress Protein
                                    ctc
ļ÷
                        (CT407)
                  0534
                                    dksA
                                              DnaK Suppressor
į.
                  0032
                         (CT341)
                                    dnaJ
                                              Heat Shock Protein J
Q
                         (CT396)
                                    dnaK
                                              Hsp-70
                  0503
                         (CT110)
                                              Hsp-60_1
                  0134
                                    groEL 1
      35
                  0777
                                    groEL_2
                         (CT604)
                                              Hsp-60 2
                  0898
                         (CT755)
                                    groEL_3
                                              Hsp-60_3
                  0135
                         (CT111)
                                    groES
                                              10KDa Chaperonin
                  0502
                         (CT395)
                                    grpE
                                              HSP-70 Cofactor
                         (CT541)
                                              FKBP-type Peptidyl-prolyl Cis-Trans Isomerase
                  0661
                                    mip
       40
               Proteases
                  0144
                         (CT113)
                                    clpB
                                              Clp Protease ATPase
                  0437
                         (CT286)
                                    clpC
                                              ClpC Protease
                         (CT431)
                                    clpP_1
                                              CLP Protease
                  0520
                  0847
                         (CT706)
                                    clpP_2
                                              CLP Protease Subunit
       45
                                              CLP Protease ATPase
                         (CT705)
                  0846
                                    clpX
                  0269
                         (CT138)
                                              Dipeptidase
                  0998
                         (CT841)
                                    ftsH
                                              ATP-dependent Zinc Protease
                  0030
                         (CT343)
                                              O-Sialoglycoprotein Endopeptidase_l-
                                    gcp_l-
                                              O-Sialoglycoprotein Endopeptidase_2
                         (CT197)
                  0194
                                    gcp_2
       50
                  0979
                         (CT823)
                                    htrA
                                              DO Serine Protease
                                              Insulinase family/Protease III
                  0957
                         (CT806)
                                    ide
                                              Lon ATP-dependent Protease
                         (CT344)
                  0027
                                    lon
                  1017
                         (CT859)
                                    lytB
                                              Metalloprotease
                         (CT851)
                                              Methionine Aminopeptidase
                                    map
       55
                                              Leucyl Aminopeptidase A
                  0385
                         (CT045)
                                    pepA
```

Oligopeptidase

6-Phosphogluconate Dehydrogenase

0360

0136

(CT112)

pepF

(CT063)

	0813	(CT574)	pepP	Aminopeptidase P
	0613	(CT494)	sohB	Protease
	0555	(CT441)	tsp	Tail-Specific Protease
	0344	(CT072)	yaeL	Metalloprotease
5	1890	(CT824)		Zinc Metalloprotease (insulinase family)
	Protein Is	omerases		
	0227	(CT176)	dsbB	Disulfide bond Oxidoreductase
	0786	(CT595)	dsbD	Thio:disulfide Interchange Protein
	0228	(CT177)	dsbG	Disulfide Bond Chaperone
10	0933	(CT783)		Predicted Disulfide Bond Isomerase
	0926	(CT780)		Thioredoxin Disulfide Isomerase

Transcription

```
RNA Degradation
                                             Polyribonucleotide Nucleotidyltransferase
                 0999 (CT842)
       5
                                             Ribonuclease III
                 0054
                        (CT297)
                                             Ribonuclease HII_1
                 0119
                        (CT029)
                                   mhB 1
                                             Ribonuclease HII 2
                        (CT008)
                 1068
                                   mhB 2
                                             Ribonuclease P Protein Component
                 0934
                        (CT784)
                                   mpA
                 0504
                        (CT397)
                                   vacB
                                             Ribonuclease Family
      10
               RNA Elongation & Termination Factors
                 0741
                       (CT636)
                                   greA
                                             Transcription Elongation Factor
                        (CT097)
                                             N Utilization Protein A
                 0316
                                   nusA
                                             Transcriptional Antitermination
                        (CT320)
                 0076
                                   nusG
                  0845
                        (CT704)
                                   pcnB_1
                                             Poly A Polymerase_1
      15
                  0966
                        (CT410)
                                   pcnB_2
                                             PolyA Polymerase_2
                  0610 (CT491)
                                             Transcription Termination Factor
                                   rho
               RNA Methylases
                  0674 (CT553)
                                             RNA Methyltransferase
                                   fmu
                  1059
                        (CT354)
                                   kgsA
                                             Dimethyladenosine Transferase
20
                  0187
                        (CT133)
                                             Predicted Methylase
                  0530
                        (CT403)
                                   spoU_I
                                             rRNA Methylase_I
                                             rRNA Methylase_2
                  0660
                        (CT540)
                                   spoU_2
                                              tRNA (Guanine N-1)-Methyltransferase
                  0117 (CT027)
                                   tmD
Ш
                                             rRNA Methyltransferse
                  0885
                        (CT742)
                                   ygcA
      25
                        (CT829)
                                   yggH
                                             Predicted rRNA Methylase
ļ.
                  0986
Ф
                  0987
                       (CT830)
                                   ytgB
                                              Predicted rRNA Methylase
M
               RNA Modification
                                             Methionyl tRNA Formyltransferase
                  0649 (CT530)
                                   fmt
                                              tRNA Pyrophosphate Transferase
                  0910
                        (CT766)
                                   miaA
      30
                                              Predicted Pseudouridine Synthase
į.
                  0719
                        (CT658)
                                   sthB
                  0219
                        (CT193)
                                              Queuine tRNA Ribosyl Transferase
-
                                              Pseudouridylate Synthase I
                  0580
                        (CT463)
                                   truA
1=
                       (CT094)
                                              tRNA Pseudouridine Synthase
                  0319
                                   truB
Û
                                              Predicted Pseudouridine Synthetase Family
                                   yceC
                        (CT106)
                  0403
4D
      35
                                              Predicted Pseudouridine Synthase
                  0864
                        (CT723)
                                   yjbC
               RNA Polymerase & Transcription Regulators
                  0586 (CT468)
                                   atoC
                                              Two-Component Regulator
                  0362
                        (CT061)
                                    rpsD
                                              Sigma-28/WhiG Family
                  0501
                        (CT394)
                                   hrcA
                                              HTH Transcriptional Repressor
                                              Sigma Regulatory Family Protein—PP2C Phosphatase (RsbW Antagonist)
      40
                  0793
                        (CT588)
                                    rbsU
                                              RNA Polymerase Alpha
                  0626
                        (CT507)
                                    фοΑ
                  0081
                        (CT315)
                                    троВ
                                              RNA Polymerase Beta
                        (CT314)
                                              RNA Polymerase Beta'
                  0082
                                    троС
                  0756
                        (CT615)
                                              RNA Polymerase Sigma-66
                                    moD
      45
                                              RNA Polymerase Sigma-54
                  0771
                        (CT609)
                                    тоМ
                  0511
                        (CT424)
                                    rsbV_1
                                              Sigma Regulatory Factor_1
                        (CT765)
                                              Sigma Factor Regulator_2
                                    rsbV_2
                                              Sigma-Regulatory-Factor-Histidine Kinase
                  0670 (CT549)
                                    rsbW:
                                              HTH Transcriptional Regulatory Protein + Receiver Doman
                  0750 (CT630)
                                    tctD
      50
                                              HTH Transcriptional Regulator
                  1069
                        (CT009)
                                    yfgA
                                                                            Translation
               Amino Acyl tRNA Synthesis
                  0892 (CT749)
                                    alaS
                                              Alanyl tRNA Synthetase
```

Arginyl tRNA Transferase

Aspartyl tRNA Synthetase

55

0570

(CT454)

0662 (CT542)

argS aspS

```
0932
                           (CT782)
                                      cysS
                                                Cysteinyl tRNA Synthetase
                                      gatA
                    0003
                           (CT003)
                                                Glu tRNA Gln Amidotransferase (A subunit)
                    0004
                           (CT004)
                                      gatB
                                                Glu tRNA Gln Amidotransferase (B Subunit)
                                      gatC
                                                Glu tRNA Gln Amidotransferase (C subunit)
                    0002
                           (CT002)
          5
                    0560
                          (CT445)
                                      gltX
                                                Glutamyl-tRNA Synthetase
                    0946
                          (CT796)
                                                Glycyl tRNA Synthetase
                                      glyQ
                    0663
                          (CT543)
                                      hisS
                                                Histidyl tRNA Synthetase
                    0109
                                                Isoleucyl-tRNA Synthetase
                          (CT019)
                                      ileS
                    0153
                          (CT209)
                                      leuS
                                                Leucyl tRNA Synthetase
        10
                    0931
                          (CT781)
                                      lysS
                                                Lysyl tRNA Synthetase
                    0122
                          (CT032)
                                      metG
                                                Methionyl-tRNA Synthetase
                                      pheS
                    0993
                          (CT836)
                                                Phenylalanyl tRNA Synthetase, Alpha
                    0594
                          (CT475)
                                      pheT
                                                Phenylalanyl tRNA Synthetase Beta
                    0500
                          (CT393)
                                      proS
                                                Prolyi tRNA Synthetase
        15
                    0870
                          (CT729)
                                      serS
                                                Seryl tRNA Synthetase_2
                          (CT581)
                    0806
                                      thrS
                                                Threonyl tRNA Synthetase
                    0802
                          (CT585)
                                      trpS
                                                Tryptophanyl tRNA Synthetase
                    0361
                          (CT062)
                                      tytS
                                                Tyrosyl tRNA Synthetase
                    0094
                          (CT302)
                                      valS
                                                Valyl tRNA Synthetase
20
                 Peptide Chain Initiation, Elongation & Termination
                          (CT353)
                    1067
                                     def
                                                Polypeptide Deformylase
                   0184
                          (CT122)
                                     efp_l
                                                Elongation Factor P_1
                   0895
                          (CT752)
                                      efp_2
                                                Elongation Factor P_2
                   0550
                          (CT437)
                                                Elongation Factor G
                                      fusA
       25
                   0073
                          (CT323)
                                     infA
                                                Initiation Factor IF-1
Þ
                   0317
                          (CT096)
                                     infB
                                                Initiation Factor-2
TO THE
                   0990
                          (CT833)
                                     infC
                                                Initiation Factor 3
                          (CT023)
                                     pfrA
                                                Peptide Chain Releasing Factor 1
2
                   0576
                          (CT459)
                                     prfB
                                                Peptide Chain Release Factor 2
-
       30
                          (CT800)
                                     pth
                   0950
                                                Peptidyl tRNA Hydrolase
į.
                   0318
                          (CT095)
                                     rbfA
                                                Ribosome Binding Factor A
įψ
                   0699
                          (CT677)
                                     пſ
                                                Ribosome Releasing Factor
1=
                   0697
                          (CT679)
                                     tsf
                                                Elongation Factor TS
Đ
                   0074
                         (CT322)
                                     tufA
                                                Elongation Factor Tu
Ø
       35
                 Ribosomal Proteins
                   0078 · (CT318)
                                     rl1
                                                L1 Ribosomal Protein
                   0644
                          (CT525)
                                     rl2
                                                L2 Ribosomal Protein
                   0647
                          (CT528)
                                     rl3
                                                L3 Ribosomal Protein
                   0646
                          (CT527)
                                     rl4
                                                L4 Ribosomal Protein
       40
                                                L5 Ribosomal Protein
                   0635
                          (CT516)
                                     rl5
                   0633
                          (CT514)
                                     rl6
                                                L6 Ribosomal Protein
                   0080
                          (CT316)
                                     rl7
                                                L7/L12 Ribosomal Protein
                   0953
                          (CT803)
                                     rl9
                                                L9 Ribosomal Protein
                   0079
                          (CT317)
                                     rllO
                                                L10 Ribosomal Protein
       45
                   0077
                          (CT319)
                                     rl11
                                               L11 Ribosomal Protein
                   0247
                          (CT125)
                                     rl13
                                               L13 Ribosomal Protein
                   0637
                          (CT518)
                                     rl14
                                               L14 Ribosomal Protein
                   0630
                          (CT511).
                                     _rl15__
                                               L15 Ribosomal Protein
                          (CT521)
                   0640
                                     rl16
                                               L16 Ribosomal Protein
       50
                   0625
                          (CT506)
                                     rl17
                                               L17 Ribosomal Protein
                          (CT513)
                                     rl18
                   0632
                                               L18 Ribosomal Protein
                   0118
                          (CT028)
                                     rl19
                                               L19 Ribosomal Protein
                   0992
                          (CT835)
                                     rl20
                                               L20 Ribosomal Protein
                   0546
                          (CT420)
                                     rl21
                                               L21 Ribosomal Protein
       55
                   0642
                         (CT523)
                                     rl22
                                               L22 Ribosomal Protein
                         (CT526)
                                     rl23
                   0645
                                               L23 Ribosomal Protein
```

```
0545
                         (CT419)
                                    r127
                                              L27 ribosomal protein
                  0327
                         (CT086)
                                    ri28
                                              L28 Ribosomal Protein
                  0639
                         (CT520)
                                    rl29
                                              L29 Ribosomal Protein
        5
                  0112
                         (CT022)
                                    rl31
                                              L31 Ribosomal Protein
                         (CT810)
                                    rl32
                  0961
                                              L32 Ribosomal Protein
                  0250
                         (CT150)
                                    rl33
                                              L33 Ribosomal Protein
                         (CT785)
                                    rl34
                                              L34 Ribosomal Protein
                  0935
                         (CT834)
                                    rl35
                                              L35 Ribosomal Protein
                  0991
      10
                         (CT786)
                                    rl36
                                              L36 Ribosomal Protein
                  0936
                  0315
                         (CT098)
                                    rs l
                                              S1 Ribosomal Protein
                         (CT680)
                                    rs2
                                              S2 Ribosomal Protein
                  0696
                        (CT522)
                                    rs3
                                              S3 Ribosomal Protein
                 0641
                                              S4 Ribosomal Protein
                  0733
                        (CT626)
                                    rs4
      15
                  0631
                        (CT512)
                                    rs5
                                              S5 Ribosomal Protein
                         (CT801)
                                              S6 Ribosomal Protein
                  0951
                                    rs6
                 0551
                        (CT438)
                                    rs7
                                              S7 Ribosomal Protein
                  0634
                        (CT515)
                                              S8 Ribosomal Protein
                                    rs8
                  0246
                         (CT126)
                                    rs9
                                              S9 Ribosomal Protein
      20
                  0549
                         (CT436)
                                    rs I O
                                              S10 Ribosomal Protein
0627
                        (CT508)
                                    rs I I
                                              S11 Ribosomal Protein
                 0552
                        (CT439)
                                    rs12
                                              S12 Ribosomal Protein
                  0628
                        (CT509)
                                    rs13
                                              S13 Ribosomal Protein
                  0937
                        (CT787)
                                    rs14
                                              S14 Ribosomal Protein
     25
                  1000
                        (CT843)
                                    rs 15
                                              S15 Ribosomal Protein
                 0116
                        (CT026)
                                    rs16
                                              S16 Ribosomal Protein
U
                 0638
                        (CT519)
                                    rs 17
                                              S17 Ribosomal Protein
I
                 0952
                        (CT802)
                                    rs18
                                              S18 Ribosomal Protein
                        (CT524)
                                              S19 Ribosomal Protein
                 0643
                                    rs 19
ļ
     30
                 0754
                        (CT617)
                                    rs20
                                              S20 Ribosomal Protein
ļф
                 0031 (CT342)
                                    rs21
                                              S21 Ribosomal Protein
ļ-L
Ø
     35
D
                                                                          Other Categories
              Chlamydia-Specific Proteins
                 0561
                        (CT446)
                                              CHLPS Euo Protein
                 0804
                        (CT583)
                                   Gp6D
                                              CHLTR Plasmid Paralog
                        (CT119)
                 0186
                                              Similarity to IncA 1
     40
                 0291
                        (CT232)
                                   incB
                                              Inclusion Membrane Protein B
                 0292
                        (CT233)
                                   incC
                                              Inclusion Membrane Protein C
                 1026
                        (CT377)
                                              LtuA Protein
                 0333
                        (CT080)
                                              LtuB Protein
                        (CT871)
                 0005
                                   pmp I
                                              Polymorphic Outer Membrane Protein G Family
     45
                 0013
                        (CT871)
                                   pmp_2
                                              Polymorphic Outer Membrane Protein G Family
                 0014
                        (CT871)
                                              Polymorphic Outer Membrane Protein G Family
                                   pmp_3
                 0015
                        (CT871)
                                              PMP 3 (frame-shift with 0014)
                                   pmp 3
                 0016
                        (CT874)
                                              Polymorphic Outer Membrane Protein G Family
                                   pmp_4
                 0017
                        (CT871)
                                   pmp_4
                                             PMP_4 (frame-shift with 0016)
     50
                 0018
                        (CT874)
                                   pmp_5
                                              Polymorphic Outer Membrane Protein G Family
                 0019
                        (CT871)
                                              PMP_5 (frame-shift with 0018)
                                   pmp_5
                 0444
                        (CT871)
                                              Polymorphic Outer Membrane Protein G/I Family
                                   pmp_6
                 0445
                        (CT871)
                                   pmp_7
                                              Polymorphic Outer Membrane Protein G Family
                 0446
                        (CT871)
                                              Polymorphic Outer Membrane Protein G Family
                                   pmp_8
     55
                        (CT871)
                                              Polymorphic Outer Membrane Protein G/I Family
                 0447
                                   pmp 9
                 0450
                        (CT871)
                                             Polymorphic Outer Membrane Protein G Family
                                   pmp_10
                        (CT871)
                                   pmp_10
                                             PMP_10 (Frame-shift with 0450)
```

0636 (CT517)

rl24

L24 Ribosomal Protein

```
Polymorphic Outer Membrane Protein (truncated) A/I Family
                  0452
                         (CT874)
                                    pmp_12
                  0453
                         (CT871)
                                    pmp_13
                                              Polymorphic Outer Membrane Protein G Family
                                              Polymorphic Outer Membrane Protein H Family
                         (CT872)
                                    pmp_14
        5
                                              Polymorphic Outer Membrane Protein E Family
                         (CT869)
                                    pmp_15
                  0466
                         (CT869)
                                              Polymorphic Outer Membrane Protein E Family
                  0467
                                    pmp_16
                  0468
                         (CT869)
                                    pmp_17
                                              Polymorphic Outer Membrane Protein E Family
                                              PMP_17 (Frame-shift with 0468)
                  0469
                         (CT869)
                                    pmp_17
                                              PMP_17 (Frame-shift with 0469)
                  0470
                         (CT869)
                                    pmp 17
      10
                         (CT870)
                                              Polymorphic Outer Membrane Protein E/F Family
                  0471
                                    pmp_18
                  0539
                         (CT412)
                                    pmp_19
                                              Polymorphic Membrane Protein A Family
                         (CT413)
                                              Polymorphic Membrane Protein B Family
                  0540
                                    pmp_20
                        (CT812)
                                              Polymorphic Membrane Protein D Family
                  0963
                                    pmp_21
                  0562
                                              CHLPS 43 kDa Protein Homolog_1
      15
                  0927
                                              CHLPS 43 kDa Protein Homolog_2
                  0928
                                              CHL?S 43 kDa Protein Homolog_3
                  0929
                                              CHL2S 43 kDa Protein Homolog 4
                  0728
                        (CT622)
                                              CHLPN 76kDa Homolog_I (CT622)
                  0729
                        (CT623)
                                              CHLPN 76kDa Homolog_2 (CT623)
      20
0133
                         (CT109)
                                              CHLPS Hypothetical Protein
0332
                        (CT081)
                                              CHL"R T2 Protein
               Miscellaneous Enzymes/Conserved Prox ins
                  0193
                                   argR
                                              Possi le Arginine Repressor
Ü
                  1046
                                              Aron atic Amino Acid Hydroxylase
      25
                  0232
-
                                              Similarity to 5'-Methylthioadenosine Nucleosidase
                  0128
                        (CT035)
m
                                              Biotin Protein Ligase
                  0513
                        (CT426)
                                              Fe-S Oxidoreductase_1
M
                  0911
                        (CT767)
                                              Fe-S Oxidoreductase 2
                  0373
                        (CT057)
                                   gcpE
                                             GcpE Protein
1-6
      30
                  0407
                        (CT103)
                                             HAD Superfamily Hydrolase/Phosphatase
ļ.
                  0917
                        (CT771)
                                             Hydrolase/Phosphatase Homolog
Ьd
                  0488
                        (CT385)
                                             HIT Family Hydrolase
                                   ycfF
ļab
                  0701
                        (CT675)
                                   karG
                                             Arginine Kinase
ū
                  0526
                        (CT399)
                                   kpsF
                                             GutQ/KpsF Family Sugar-P Isomerase
      35
                  0919
                        (CT773)
                                             Leucine Dehydrogenase
                                   ldh
                  0022
                        (CT349)
                                   maf
                                             Maf protein
                 0997
                        (CT840)
                                   mesJ
                                             PP-loop superfamily ATPase
                  0151
                        (CT148)
                                   mhpA
                                             Monooxygenase
                 0730
                        (CT624)
                                             Integral Membrane Protein
      40
                 0861
                        (CT720)
                                             NifU-Related Protein
                 0479
                        (CT380)
                                   phnP
                                             Metal Dependent Hydrolase
                 0106
                        (CT015)
                                   phoH
                                             ATPase
                 0329
                        (CT084)
                                             Phopholipase D Superfamily
                        (CT284)
                 0435
                                             Phospholipase D Superfamily
      45
                 0581
                        (CT464)
                                             Phosphoglycolate Phosphatase
                 0897
                        (CT754)
                                             Predicted Phosphohydrolase
                 0509
                        (CT422)
                                             Predicted Metalloenzyme
                        (CT375)
                  1030
                                             Predicted D-Amino Acid Dehyrogenase
                                             SAM Dependent Methyltransferase
                 0531
                        (CT404)
     50
                 0337
                        (CT076)
                                   smpB
                                             Small Protein B
                 0394
                        (CT256)
                                   tlyC_l
                                             CBS Domain Protein (Hemolysin Homolog)_l
                 0510
                        (CT423)
                                   tlyC_2
                                             CBS Domains (Hemolysin Homolog)_2
                 0382
                        (CT048)
                                   yabC
                                             SAM-Dependent Methytransferase
                 0787
                        (CT594)
                                   yabD
                                             PHP Superfamily (Urease/Pyrimidinase) Hydrolase
     55
                 0611
                        (CT492)
                                   yacE
                                             Predicted Phosphatase/Kinase
                                             Sugar Nucleotide Phosphorylase
                 0579
                        (CT462)
                                   yacM
                 0578
                       (CT461)
                                   yael
                                             Phosphohydrolase
```

Polymorphic Outer Membrane Protein G Family

0451

(CT871)

pmp 11

```
CT071 Hypothetical Protein
                  0345
                         (CT071)
                                    yaeM
                                               YaeS family Hypothetical Protein
                  0566
                         (CT450)
                                    yaeS
                         (CT472)
                                               YagE family
                  0591
                                    yagE
                         (CT335)
                                    ybaB
                                               YbaB family Hypothetical Protein
                  0039
        5
                                    ybbP
                                               YbbP family Hypothetical Protein
                         (CT012)
                  0101
                                               iojap Superfamily Ortholog
                  0915
                         (CT769)
                                    ybeB
                  0137
                         (CT108)
                                    ybgi
                                               ACR family
                         (CT402)
                                              ATPase
                  0529
                                    ycaH
                         (CT287)
                                    ycbF
                                              PP-loop Superfamily ATPase
                  0438
      10
                         (CT627)
                                              YceA Hypothetical Protein
                  0734
                                    yceA
                                    ychB
                  0954
                         (CT804)
                                               Predicted Kinase
                  0261
                         (CT217)
                                    ydaO
                                              PP-Loop Superfamily ATPase
                                    ydhO
                                              Polysaccharide Hydrolase-Invasin Repeat Family
                  0245
                         (CT127)
                  0573
                         (CT457)
                                    yebC
                                               YebC Family Hypothetical Protein
      15
                  0689
                         (CT687)
                                    yfhO_1
                                              NifS-related Aminotransferase_1
                  0862
                         (CT721)
                                    yfhO_2
                                              NifS-related Aminotransferase_2
                  0547
                        (CT434)
                                              YgbB Family Hypothetical Protein
                                    ygbB
                         (CT184)
                  0237
                                    yggF
                                               YggF Family Hypothetical Protein
                  0775
                         (CT606)
                                    yggV
                                               YggV Family Hypothetical Protein
      20
                  0396
                         (CT258)
                                    yhfO_3
                                              NifS-related Aminotransferase_3
0605
                         (CT487)
                                    yhhF
                                              Predicted Methylase
                  0575
                         (CT458)
                                    yhhY
                                              Amino Group Acetyl Transferase
                  0592
                         (CT473)
                                    yidD
                                              YidD Family
                  0982
                         (CT825)
                                    yigN
                                              YigN Family Hypothetical Protein
      25
                  0657
                        (CT537)
                                    yjeE
                                              YjeE Hypothetical Protein
                                              Yohl Predicted Oxidoreductase
                  0768
                         (CT644)
                                    yohl
ij
                  0336
                        (CT077)
                                    yojL
                                              YojL Hypothetical Protein
H
                  0217
                         (CT140)
                                    ypdP
                                              YpdP Hypothetical Protein
                         (CT212)
                  0140
                                    yqdE
                                              YqdE Hypothetical Protein
<u>}-</u>
      30
                         (CT221)
                                    yqfU
                  0263
                                              YqfU Hypothetical Protein
÷
                 0139
                        (CT211)
                                    yqgE
                                              YqgE Hypothetical Protein
0270
                         (CT137)
                                    ywlC
                                              SuA5 Superfamily-related Protein
-
                  0879
                         (CT738)
                                              Metal Dependent Hydrolase
                                    yycJ
4
     35
                                                            Homologs to CHLTR Hypothetical Coding Genes
                        (CT001)
                 0001
                                    CT001 Hypothetical Protein
                  0020
                         (CT351)
                                    CT351 Hypothetical Protein
                        (CT350)
                  0021
                                    CT350 Hypothetical Protein
                        (CT345)
                  0026
                                    CT345 Hypothetical Protein
     40
                        (CT339)
                 0035
                                    CT339 Hypothetical Protein
                 0036
                        (CT338)
                                    CT338 Hypothetical Protein
                 0055
                        (CT296)
                                    CT296 Hypothetical Protein
                 0062
                        (CT289)
                                    CT289 Hypothetical Protein
                        (CT288)
                 0065
                                    CT288 Hypothetical Protein
     45
                 0068
                        (CT360)
                                    CT360 Hypothetical Protein
                 0071
                        (CT325)
                                    CT325 Hypothetical Protein
                        (CT324)
                 0072
                                    CT324 Hypothetical Protein
                        (CT311)
                 0085
                                    CT311 Hypothetical Protein
                 0087
                        (CT309)
                                    CT309 Hypothetical Protein
     50
                 0093
                        (CT303)
                                    CT303 Hypothetical Protein
                 0100
                        (CT011)
                                    CT011 Hypothetical Protein
                 0104
                        (CT017)
                                    CT017 Hypothetical Protein
                 0105
                        (CT016)
                                    CT016 Hypothetical Protein
                        (CT058)
                                    CT058 Hypothetical Protein_1
                 0107
     55
                        (CT018)
                 0108
                                    CT018 Similarity
                 0111
                        (CT021)
                                    CT021 Hypothetical Protein
                        (CT031)
                                    CT031 Hypothetical Protein
                 0121
```

```
0129
                         (CT036)
                                    CT036 Similarity
                         (CT114)
                                     CT114 Hypothetical Protein
                   0150
                         (CT147)
                                     CT147 Hypothetical Protein
                         (CT149)
                                    CT149 Hypothetical Protein
                   0152
        5
                         (CT153)
                                    CT153 Hypothetical Protein
                   0176
                                    CT132 Hypothetical Protein
                   0188
                         (CT132)
                                    CT131 Hypothetical Protein
                   0189
                         (CT131)
                                     CT203 Hypothetical Protein
                   0206
                         (CT203)
                                     CT178 Hypothetical Protein
                   0229
                         (CT178)
                                     CT179 Hypothetical Protein
       10
                   0230
                         (CT179)
                                     CT181 Hypothetical Protein
                   0234
                         (CT181)
                                     CT151 Hypothetical Protein
                   0249
                         (CT151)
                                     CT144 Hypothetical Protein_1
                         (CT144)
                   0253
                                     CT143 Hypothetical Protein_I
                   0254
                         (CT143)
       15
                                     CT142 Hypothetical Protein_1
                   0255
                         (CT142)
                         (CT144)
                                     CT144 Hypothetical Protein_2
                   0256
                                     CT143 Hypothetical Protein_2
                         (CT143)
                   0257
                                     CT142 Hypothetical Protein_2
                         (CT142)
                   0259
                                     CT191 Hypothetical Protein
                         (CT191)
                   0276
       20
                                     CT195 Hypothetical Protein
                   0288
                         (CT195)
(CT234)
                                     CT234 Hypothetical Protein
                   0293
                                     CT368 Hypothetical Protein
                         (CT242)
                   0301
                                     CT244 Hypothetical Protein
                   0303
                         (CT244)
                                     CT249 Similarity
                   0308
                          (CT249)
       25
                   0312
                          (CT101)
                                     CT101 Hypothetical Protein
                                     CT085 Hypothetical Protein
                   0328
                          (CT085)
.
WN
                                     CT083 Hypothetical Protein
                   0330
                          (CT083)
                                     CT082 Hypothetical Protein
                          (CT082)
                   0331
                                     CT079 Similarity
                   0334
                          (CT079)
30
                   0342
                          (CT073)
                                     CT073 Hypothetical Protein
ļф
                                     (frame-shift with 0342?)
                   0343
                          (CT073)
14
                                     CT066 Hypothetical Protein
                   0350
                          (CT066)
CT058 Hypothetical Protein_2
                          (CT058)
                   0369
J
                                     CT058 Hypothetical Protein_3
                   0370
                          (CT058)
Ü
       35
                          (CT056)
                                     CT056 Hypothetical Protein
                   0379
                          (CT053)
                                     CT053 Hypothetical Protein
                   0381
                          (CT326)
                                     CT326 Similarity
                          (CT047)
                                     CT047 Hypothetical Protein
                   0383
                                     CT043 Hypothetical Protein
                   0387
                          (CT043)
       40
                   0389
                          (CT041)
                                      CT041 Hypothetical Protein
                                      CT038 Hypothetical Protein
                   0393
                          (CT038)
                          (CT257)
                                     CT257 Hypothetical Protein
                   0395
                          (CT253)
                                      CT253 Hypothetical Protein
                   0399
                   0400
                          (CT254)
                                      CT254 Hypothetical Protein
        45
                   0401
                          (CT255)
                                      CT255 Hypothetical Protein
                                      CT105 Hypothetical Protein
                   0405
                          (CT105)
                                      CT102 Hypothetical Protein
                    0408
                          (CT102)
                                      CT260 Hypothetical Protein
                    0409
                          (CT260)
                          (CT262)
                                      CT262 Hypothetical Protein
                    0411
        50
                    0412
                          (CT263)
                                      CT263 Hypothetical Protein
                                      CT266 Hypothetical Protein
                          (CT266)
                    0415
                          (CT271)
                                      CT271 Hypothetical Protein
                    0420
                    0422
                          (CT273)
                                      CT273 Hypothetical Protein
                                      CT274 Hypothetical Protein
                    0423
                          (CT274)
        55
                                      CT276 Hypothetical Proteins
                    0425
                          (CT276)
                    0426
                          (CT277)
                                      CT277 Similarity
                                      CT283 Hypothetical Protein
                    0434
                         (CT283)
```

		0441	(CT007)	CT007 Hypothetical Protein
		0442	(CT006)	CT006 Hypothetical Protein
		0443	(CT005)	CT005 Hypothetical Protein
		0474	(CT365)	CT365 Hypothetical Protein
	• 5	0476	(CT865)	CT865 Hypothetical Protein
		0480	(CT383)	CT383 Hypothetical Protein
		0485	(CT382)	CT382.1 Hypothetical Protein
		0487	(CT384)	CT384 Hypothetical Protein
		0489	(CT386)	CT386 Hypothetical Protein
	10	0490	(CT387)	CT387 Hypothetical Protein
		0491	(CT389)	CT389 Hypothetical Protein
		0496	(CT391)	CT391 Hypothetical Protein
,		0497	(CT388)	CT388 Hypothetical Protein
		0506	(CT421)	CT421 Hypothetical Protein
	15	0507	(CT421)	CT421.1 Hypothetical Protein
	13	0508		, - ·
			(CT421)	CT421.2 Hypothetical Protein
		0512	(CT425)	CT425 Hypothetical Protein
		0514	(CT427)	CT427 Hypothetical Protein
.eeq	20	0518	(CT429)	CT429 Hypothetical Protein
<u>.</u>	20	0522	(CT433)	CT433 Hypothetical Protein
\L		0525	(CT398)	CT398 Hypothetical Protein
2 <u>1</u> 222 2226		0533	(CT406)	CT406 Hypothetical Protein
أببا		0537	(CT814)	CT814.1 Hypothetical Protein
Ü		0538	(CT814)	CT814 Hypothetical Protein
	25	0554	(CT440)	CT440 Hypothetical Protein
		0559	(CT441)	CT441.1 Hypothetical Protein
m		0565	(CT449)	CT449 Hypothetical Protein
5		0572	(CT456)	CT456 Hypothetical Protein
ļ.		0582	(CT465)	CT465 Hypothetical Protein
	30	0583	(CT466)	CT466 Hypothetical Protein
-		0588	(CT469)	CT469 Hypothetical Protein
ļ.		0589	(CT470)	CT470 Hypothetical Protein
ļ=		0590	(CT471)	CT47! Hypothetical Protein
ū		0593	(CT474)	CT474 Hypothetical Protein
١Ī	35	0595	(CT476)	CT476 Hypothetical Protein
		0601	(CT483)	CT483 Hypothetical Protein
		0602	(CT484)	CT484 Hypothetical Protein
		0606	(CT488)	CT488 Hypothetical Protein
		0609	(CT490)	•
	40			CT490 Hypothetical Protein
	70	0622	(CT503)	CT503 Hypothetical Protein
		0623	(CT504)	CT504 Hypothetical Protein
		0648	(CT529)	CT529 Hypothetical Protein
	•	0658	(CT538)	CT538 Hypothetical Protein
	15	0667	(CT546)	CT546 Hypothetical Protein
	45	0668	(CT547)	CT547 Hypothetical Protein
		0669	(CT548)	CT548 Hypothetical Protein
		0671	(CT550)	CT550 Hypothetical Protein
		0673	(CT552)	CT552 Hypothetical Protein
-		0675	(CT696)	CT696 Hypothetical Protein
	50	0676	(CT695)	CT695 Similarity
		0681	(CT691)	CT691 Hypothetical Protein
		0687	(CT482)	CT482 Hypothetical Protein
		0688	(CT481)	CT481 Hypothetical Protein
		0700	(CT676)	CT676 Hypothetical Protein
	55	0705	(CT671)	CT671 Hypothetical Protein
		0703		•
			(CT670)	CT670 Hypothetical Protein
		0708	(CT668)	CT668 Hypothetical Protein

```
0709
                                               (CT667)
                                                                    CT667 Hypothetical Protein
                                 0710
                                               (CT666)
                                                                    CT666 Hypothetical Protein
                                                                    CT665 Hypothetical Protein
                                 0711
                                               (CT665)
                                                                    CT663 Hypothetical Protein
                                              (CT663)
                                 0713
              5
                                                                    CT656 Hypothetical Protein
                                 0717
                                              (CT656)
                                 0718
                                              (CT657)
                                                                    CT657 Hypothetical Protein
                                                                    CT659 Hypothetical Protein
                                 0720
                                              (CT659)
                                                                    CT654 Hypothetical Protein
                                 0722
                                              (CT654)
                                              (CT652)
                                                                    CT652.1 Hypothetical Protein
                                 0725
           10
                                 0726
                                              (CT620)
                                                                    CT620 Hypothetical Protein
                                                                    CT619 Hypothetical Protein
                                 0727
                                              (CT619)
                                 0739
                                              (CT638)
                                                                    CT368 Hypothetical Protein
                                 0742
                                              (CT635)
                                                                    CT635 Hypothetical Protein
                                 0746
                                              (CT632)
                                                                    CT632 Hypothetical Protein
           15
                                 0747
                                              (CT631)
                                                                    CT631 Hypothetical Protein
                                 0751
                                              (CT651)
                                                                    CT651 Hypothetical Protein
                                 0755
                                              (CT616)
                                                                    CT616 Hypothetical Protein
                                0760
                                              (CT611)
                                                                    CT611 Hypotheti:al Protein
                                              (CT610)
                                                                    CT610 Hypotheti:al Protein
                                 0761
The first the first two of the second second
         20
                                              (CT648)
                                                                    CT648 Hypotheti:al Protein
                                0764
                                0765
                                              (CT647)
                                                                    CT647 Hypotheti:al Protein
                                0766
                                              (CT646)
                                                                   CT646 Hypothetical Protein
                                0767
                                              (CT645)
                                                                   CT645 Hypotheti- al Protein
                                0770
                                              (CT642)
                                                                   CT642 Hypotheti :al Protein
          25
                                              (CT606)
                                0774
                                                                   CT606.1 Hypothetical Protein
                                0776
                                              (CT605)
                                                                   CT605 Hypothetical Protein
                                0779
                                              (CT602)
                                                                   CT602 Hypothetical Protein
                                0783
                                              (CT598)
                                                                   CT598 Hypothetical Protein
7
                                0791
                                              (CT590)
                                                                   CT590 Hypothetical Protein
ļ-
          30
                                             (CT589)
                                0792
                                                                   CT589 Hypothetical Protein
H
                                0803
                                              (CT584)
                                                                   CT584 Hypothetical Protein
Į.
                                0807
                                             (CT580)
                                                                   CT580 Hypothetical Protein
14
                                0808
                                             (CT579)
                                                                   CT579 Hypothetical Protein
Ð
                                             (CT578)
                                0809
                                                                   CT578 Hypothetical Protein
₩ 35
                                0810
                                              (CT577)
                                                                   CT577 Hypothetical Protein
                                0814
                                              (CT573)
                                                                   CT573 Hypothetical Protein
                                                                   CT569 Hypothetical Protein
                                0818
                                             (CT569)
                                0819
                                             (CT568)
                                                                   CT568 Hypothetical Protein
                                0820
                                             (CT567)
                                                                   CT567 Hypothetical Protein
          40
                                0821
                                             (CT566)
                                                                   CT566 Hypothetical Protein
                                0822
                                             (CT565)
                                                                   CT565 Hypothetical Protein
                                0827
                                             (CT560)
                                                                   CT560 Hypothetical Protein
                                0834
                                             (CT556)
                                                                   CT556 Hypothetical Protein
                                             (CT700)
                                0840
                                                                   CT700 Hypothetical Protein
          45
                                0842
                                             (CT702)
                                                                   CT702 Hypothetical Protein
                                             (CT702)
                                0843
                                                                   CT702 Hypothetical Protein
                                0852
                                             (CT711)
                                                                   CT711 Hypothetical Protein
                                0853
                                             (CT712)
                                                                   CT712 Hypothetical Protein
                                0857
                                             (CT716)
                                                                   CT716 Hypothetical Protein
          50
                                0859
                                             (CT718)
                                                                   CT718 Hypothetical Protein
                                0865
                                             (CT724)
                                                                   CT724 Hypothetical Protein
                                0869
                                             (CT728)
                                                                   CT728 Hypothetical Protein
                                0874
                                             (CT733)
                                                                   CT733 Hypothetical Protein
                                0875
                                             (CT734)
                                                                   CT734 Hypothetical Protein
         55
                                0884
                                             (CT741)
                                                                   CT741 Hypothetical Protein
                                0887
                                             (CT744)
                                                                   CHLTR Possible Phosphoprotein
                                0896
                                            (CT753)
                                                                  CT753 Hypothetical Protein
```

		0906	(CT763)	CT763	Hypothetical Protein	
		0908	(CT764)	CT764	Hypothetical Protein	
		0912	(CT768)	CT768	Hypothetical Protein	
		0925	(CT779)	CT779	Hypothetical Protein	
	5	0938	(CT788)	CT788	Hypothetical Protein	
		0939	(CT790)	CT790 I	Hypothetical Protein	
		0943	(CT794)	CT794.	1 Hypothetical Protein	
		0945	(CT795)	CT795 I	Hypothetical Protein	
		0956	(CT805)	CT805 I	Hypothetical Protein	•
	10	0960	(CT809)	CT809 1	Hypothetical Protein	
		0989	(CT832)	CT832 I	Hypothetical Protein	•
		0994	(CT837)	CT837 I	Hypothetical Protein	
		0995	(CT838)	CT838 I	Hypothetical Protein	
		0996	(CT839)	CT839 I	Hypothetical Protein	
	15	1002	(CT845)		Hypothetical Protein	
		1003	(CT846)		Hypothetical Protein	
		1004	(CT847)	_	Hypothetical Protein	
		1005	(CT848)		Hypothetical Protein	
		1006	(CT849)		Hypothetical Protein	
,FESTA	20	1007	(CT849)		1 Hypothetical Protein	i.
		1008	(CT850)		Hypothetical Protein	
1		1010	(CT852)		Hypothetical Protein	j.
2		1011	(CT853)		Hypothetical Protein	į.
11.00		1015	(CT857)		Hypothetical Protein	I .
· (I)	25	1016	(CT858)		Hypothetical Protein	<u>:</u>
gent destruction of the second second of the second		1019	(CT860)		Hypothetical Protein	j .
Ø		1020	(CT861)		Hypothetical Protein	i.
Man,		1022	(CT863)		Hypothetical Protein	
ä		1032	(CT373)			
	30	1032	(CT373)		Hypothetical Protein	
į.	50	1033			Hypothetical Protein	
i in		1057	(CT371)		Hypothetical Protein	
		1058	(CT356) (CT355)		Hypothetical Protein	
.Fi		1061	(CT330)		Aypothetical Protein	
	35	1073	(CT371)		Hypothetical Protein Hypothetical Protein	
7 East	73	1075	(013/1)	C13/11	Typothetical Protein	
					Coding Genes Not in C. trachomatis	
		0486			Hypothetical Proline Permease	
		0279			Possible ABC Transporter Permease Protein	
	40	0505			3-Methyladenine DNA Glycosylase	
		0193		argR	Similarity to Arginine Repressor	
		1041		bioA	Adenosylmethionine-8-Amino-7-Oxononanoate Aminotransferase	
		1044		bioB	Biotin Synthase	
		1042		bioD	Dethiobiotin synthetase	
	45	0585		OIOD	Similarity to Cps IncA_2	
		0562			CHLPS 43 kDa Protein Homolog_1	
		0927			CHLPS 43 kDa Protein Homolog_2	
		0928			CHLPS 43 kDa Protein Homolog_3	
	= =====	0929		 -	CHLPS 43 kDa Protein Homolog_4	
	50	1045			Conserved Hypothetical Membrane Protein	
		0251			Conserved Hypothetical Protein	
		0278			Conserved Outer Membrane Lipoprotein Protein	
		0907			CutA-like Periplasmic Divalent Cation Tolerance Protein	
		0171	•	guaA	GMP Synthase	
	55	0171				
		0608		guaB	Inosine 5'-Monophosphase Dehydrogenase	
					Uridine 5'-Monophosphate Synthase	
		0735			Uridine Kinase	

		0980		Similar to	Saccharomyces cerevisi	ae 52.9KDa Prote	ein			
		0232		Similarity	to 5'-Methylthioadenosi	ne Nucleosidase				
		1046		Tryptopha	in Hydroxylase					
		0477	yqeV	Bs Conserved	l Hypothetical Protein					
	5	0048	yqfF-l		i Hypothetical IM Protei	n				
		0587			l Hypothetical Protein					
		0143	yxjG_	_	ved Hypothetical Proteir					
		0448	yxjG_		ved Hypothetical Proteir					
		0006	0180	0440	0977	•				
	10	0007	0181							
	10		0190 -	0455	0978	•				
		0008		0456	1018					
_		0009	0203	0457	1023					
		0010	0204	0458	1027					
	15	0011	0205	0459	1029					
	13	0012	0209	0460	1040					
		0028	0210	0461	1051					
		0029	0211	0462	1052					
		0034	0212	0463	1053					
	20	0041	0213	0464	1054					1
4-	_ 20	0042	0214	0465	1055					
4Í		0043	0215	0472	1056		•			
5	25 25	0044	0216	0473	1064					
13	J.	0045	0218	0481	1065					
f)	1	0046	0220	0483	1066				•	
5 m	25	0047	0221	0492	1070					
The state of the section of the sect	ng.	0049	0222	0493	1071					
10	## ##	0050	0223	0494	1072					
		0051	0224	0498						5
And the second s		0063	0225	0499		-				
25	∮ 30	0064	0226	0516						
P		0066	0233	0517						
ž.	¥	0067	0240	0523					-	
ŀ	b.	0069	0241	0524						
¥Ĺ	,	0070	0242	0553						
4	35	0099	0243	0574						
		0124	0266	0600			•			
		0125	0267	0656						
		0126	0268	0664						
		0130	0277	0677						
	40	0131	0283	0678						
		0132	0284	0685						
		0142	0285	0686						
		0146	0287	0724						
		0147	0352	0731						
	45	0155	0353	0745						
		0156	0354	0753						
		0157	0355	0794						
		0158	0356	0795						
		0159	0357					 		
	50	0162	0358	0797						
		0163	0365	0798						
		0164	0366	0799						
		0165	0367	0829						
		0166	0368	0830						
	55	0167	0371	0831	·					
		0168	0372	0881						
		0169	0375	0882						

0170	0376	0913
0173	0391	0914
0174	0398	0930
0175	0404	0944
0177	0431	0964
0178	0432	0975
0170	0420	0076

Table 3

Chlamydia pneumoniae Genome Encoded Pr

SKRYAKSVVLGGOS

CTROL hypothetical protein KRUKDEIKYTSURRKAMLOKIIRGLSSLIVILCALNYGLIGITHNKLNIIAKLCGGVSTP ATQITYIIIGIAGVICLLSFCPFCSKKSRHSHGDSCSSGCHSHHSDKN

CPn 0002 CENTENDAM AND STATE AND STATE OF THE STATE O

889 CPH_0003 889 2370
gata-Glu ERNA Gln Amidotransferae
kinyrkysalelakavilgeltatgytohffhrieeaegovgafislckeoaleoaelidk
krsrgeploklagovpygikdnihvtglkticasrvlenyoppfdatvverikkedgilig
klimdepfamostitysafhftnipmblsrvegossogsaaavsarfcevalgsdtgssir
Opaafcgvvgfkpsygavsryglvafasslodigplantvedvalmddyfsgrdpkdats
REFFRDSFMSKLStevpkvigyprtflegleddirenffsslaifegogthlydveldil
SHAVSIYYILASAEARINLARFGGVRYGYRSPOAHTISOLVDLSRGDSFREVMRRILIG
NYVLSAERQNVYYKKATAVRAKIVKAFRTAFEKCEILAMPVCSSPAFEIGEILDPVTLYL
ODITTOAMBLAYLPATAVPSGFSKEGLPLGLQIIGQOGDOOVCOVGYSFQEHAQIKOLF

CPn_0004 2334 3833
gatB-(Petl12) Glu trnA Gln Amidotransferase (B Subunit)
EICOKCCSRRSIMSAUYADMESVIGLEVHVELNTASKLFSSALNRFGDEPNTNISTVCTG
LPGSLPVLNQSAVEKAVLFGCAVEGEISLLSRFDRKSYFYPDSPRNFQITOFEHPITRGG
RIKAIVQGEERYFELAQTHIEDDAGMLKHFGEFAGVDYNRAGVPLIEIVSKPCMFCPBGV
AVAYATSLVSLLDYIGISCONMESGSIRFDVNVSVRPKGSPELRNKVEINDANSFAFMAQA
LEAEKQRQIDEVLNOPNKDPKLVIPAATYRWDPEKKKTVLMRLKESAEDYKYFPEPDLPT
LQLTESYIERIRKTLPELPYDKYHRYIQEYGLSEDIASILISKNITAFFEVACKDCKNF
RSLSKMVTVEFGGRCKTLGVKLPSSGIFPEGVAQLVNAIDQGVITGKIAKEIADLMMESP
GKNPEEILKEKPELLPMSDEGELQKIIAEVVLANPESIVDYKNGKTKALGFLVGQIMKRT
AGKAPPKRVNELLLLELDKG

CPN_005

CPN_005

A097

6892

pp_1-Polymorphic Outer Membrane Protein

SDIHFDLGTKMRFSLCGFPLVFSFTLLSVFDTSLSATTISLTPEDSFHGDSQNAERSYNV

OAGDVY SLTGDVSISNYDDSAJNKAC FNVTSGSVTF AGNHHGLYFNNISSGTTKEGAVLC

COPPOATARFSGFSTLSF [OSPODIKEQGCLYSKNALM-LLNNYVVRFEQNGSKTKGGAIS

GANAFIVGNYDSVSFYQNAATFGGAIHSSGPLQIAVNQAEIRFAQNTAKNGSGGALYSDG

DIDHDQNAYVLFRENEALTTAIGKGGAVCCLPPSGSSTPVPLVTFSDNKQLVFERNHSIM

SGGÄTYARKLSISSGPTLFINNISYANSQNLGGAIAIDTGGEISLSAEKGT ITFQGNRT

SLEFELNGIHLLQNAKFLKLQARNGYSIEFTYDFITSEAGSTQLUINGDFKNKEYTGTILF

SGEÄSLANDPEDFEKSTIPONVLISAGYLVIKEGAEVTVSKTTQSFGSHLVLDLGTKLIAS

KEDIAITGLAIDIDSLSSSSTAAVIKANTANKQISYTDSIELISFTGNAYEDLAMRNSQT

FPLÄSLEPGAGGSVTVTAGDFLVSPHYGPGGMKLAMTGTGNKVGEFFWKKINYKRPE

KEĞÄLVENILWGNAVDVRSLMQVQETHASSLQTDRGLWIDGIGNFFHVSASEDNIRYRHN

SGĞYVLSVNNEITPKHYTSMAFSQLSSKDYAVSNNEYRHYLGSYLVQYTTSLGNIFRY

ASRMENNVNGLISRRFLQNPLMIFHFLCAYGHATNDMTDVANFFMVNSSRNNCWAIEC

GSWPLLVFENGRLFQGAIFFMKLDLVYAYQGDFKETTADGRRFSNGSLTSISVPLGIRF

EKLÄLSQDVLYDFSFSYIPDIFRKDPSCEAALVISGDSWLVPAAHVSRHAFVGSGTGRYH

FNDFTEELLCRGSIECRPHARNYNINGGSKFRF FNDYTELLCRGSIECRPHARNYNINCGSKFRF

7299 CPn_0006 No robust homolog present in Genebank/EMBL as of 11/7/98 KOEQEPLRSALLERLSEWLVLLGVPSPETTRSTPEKDANQLPKDSRNRTLESL

OPA-0007 7488 10496

No Tobust homolog present in Genebank/EMBL as of 11/7/98

KSEMPINLSLIFSFLVVIPLTDSTTSSLSTSLLDEGNPQSMRKLRILAIVLIALSIILIAG

GVVLLTVAIPGLSSVISSPAGMGACALGCVMLALGIDVLLKKREVPIVLASVTTTPGTGS

PRÉGESISGADSTIRSLPTYLLDEGHPQSMRKLRILAIVLIVFSIILIASGVVLLTVAIPG

GLSSVISSPAGMGACALGCVMLALGIDVLLKKREVPIVLASVTTTPGTGSFRSGISISGA

DSTÉLBSLPTYPLDEGHPQSMRKLRILAIVLIVFSIILIASGVVLLTVAIPGLSSIISPA

EMGAGALGCVMLAUGIDVLLKKREVPIVVPAPPIPEEVVIDDIDESSIRLQDEAPALARL

PEEMSÄREGYIKVVESHLENMKSLPYDGHGLEEKTKHDIRVVRSSLKAMVPEFLDIRRIF

EEEEPFFLSARKRLIDLATTLVERKILTEQLERNNLRKAFSVLYQDSIFKKIILNFEKLA

MKFMILSKSICRFTIIFENHEHGVANSLLHKNAVLLEKVIYRSLKAMVPEFLDIRRIF

EEEEPFFLSARKRLIDLATTLVERKILTEQLERNNLRKAFSVLYQDSIFKKIILNFEKLA

MKFMILSKSICRFTIIFENHEHGVANSLLHKNAVLLEKVIYRSLQKSVRDIGMSSAMMKI

LHGNPFFSLEDNKKTIMKEHAEMLESLSSYRVFLALSDENVVDTPSDPKWPLSGIFCR

DALSEISRDEQWQKKAHLKHQESLYTOARDRLTDQSSKENQKELEKADECY FSSMERVKK

PEIERVOERIRAIQKLYPNILEREETTGQETVTPTVQGTTASSDLTDILQRIEVSSRED

NQNQESCVKVLRSHEVEMSWEVKQEVGPKKKEPQDQMGSLERFFTEHTELEVLOKVDSK

HLSYFKKVNNKKEVQVAKFRLKVLESDLGGILAGTGSAESLLTDGELFILATRGALEKAV

FKGSLCCALASKARPYFEEDPRFQDSDTOLRALTLRLDEAKASLEEEIKRFSNLENDIAE

ERRLLKESKQTFERAGLGVLREIAVESTYDLRSLNTTWEGTPESEKVYFSMYLNYYNEEK

RRAKTRLVEMTORYTOPKYNDEKMALEAMQFNEEALLQEELSIQAPSE RRAKTRLVEMTORYRDFKMALEAMOFNEEALLQEELSIQAPSE

10780 11685 No robust homolog present in Genebank/EMBL as of 11/7/98 CKYSYLLNYPPPPRRSLGVSCSKLRSLSITLLVLGVLLLTLGIFGLTAGISFGAGLGFSA CKTSYLLINTPPPRINGLOVSCALESIES I LLUVUVILLE I SSOELIROVYCASVOVILLE SADALKISVIVEVIDEVINTVIVILLE I SSOELIROVYCASVOVILLE SADALKISVIVEVIDEVINTVIVILLE I LLUVUVILLE I SSOELIROVILLE I SSOELIROVILLE SAVANI LLUVUVILLE I SSOELIROVILLE SVOVILLE I STORI STORI LLUVUVILLE I SSOELIROVILLE SVOVILLE I STORI STOR

CPn_0009 11689 13119
No robust homolog present in Genebany/EMBL as of 11/7/98
-fdrsahaborerding:medeketiewvosidectdisfyrkschwedryadkfelrekeekmerheldeinatwirkaschayakakaafekersnehörevkovekwissglaefenoeser
Arerlreldytypevsveervlerortkkynleny/kolekkyhlkovabehyaksenke
Paeyreriekvisaeevskelorleboletwskutkaeesyfemkfdateklonkvisd
vynneletisedaemiffreet emterweeldylemkhytekaseloynsckemlakvepo
ckesptyrsjoerierikololipatynogeriosidleskverordilarekomiffevog
liffingellaviaelftoarlolvatypymefyloynfikrekveromakteryreiro
afortwekebliaaedtilkeedywilrodwilrodwilrokkekveromakteryreiro CPn 0009 11689 13119

125 13324 No robust homolog present in Genebank/EMBL as of 11/7/98

RVLAITELVEGMULLISGALFUTLGIPGUSAAIS CKYFYLRSYPPPPQHSVGS1 FGLGIGLSALGGVLMISGLL CKYFYLRSYPPPPOHSVOSI RVLA ITFLVFOMLLLISGALFITTGI FOLSAAIS
FGIGIGIGALGOVIMISCIL REIPTVRPEEI PEGVSLAPSEEPALQAAQKTLAQL
PKELDQLDTDIQEVFACLRKLADSKYESRSFLNDAKKELRVFDFV/EDTLSEIFELRQIV
AQEOMOLNFLINGGRSIMMTAESESLDLFHVSKRLGYLPSGDVRGEGLKKSAKEIVARLM
SLHCEIHKVAVAFDRNSYAMAEKAFAKALGALEFSVYRSLTQSYRDKFLESERAKIPMNG
HITWLRDDAKSCCAEKKLGMPRNVGRNLGKQSFG

CPn_0010.1 14 1574/ 14268 ELIMO BIET WEEL-LINE ENDENNEEDEN ENDENNEEDEN ERMEETE ENDELEGENE KEAPDALEERUNGERSCHEIDEN ELEGEN ENDELEGENEEN ENDELEGEN ENDELEGENEEN ENDELEGEN ENDELEGEN ENDELEGEN ENDELEGEN ENDELEGEN ENDELEGEN ENDELEGEN EL SEELEGEN EN SEELEG DKARFLFNREDHS

CPn_0011 1587 16614

gatB-(Pet112) Glu trvA Gln Amidotransferase (B Subunit)

fWYSIMTAAPAILHVSPTP9EETKFVIFKDSKSRALGITLLVVGILLVVCGAIVLSGVIS

GLSALIVCGLGISTISLGVUFLVLGILLLLRREETLEQIEAKGIAETFAADELKELEMYI

QSTEKSLEKIEGSRYSDOFFLNRATQKILDLESSLSSITSEFRDLRQLFDEEKIELLSGE

RLLEFIAANLFKQGRDVYLNLGNLADIRAYMGPNNYKVAMVIEKAKAVVHEFIVLTTMAR

FLEFFF

CPn_0012 16596 18212
gatB-(Pet112) flu tRNA Gln Amidotransferase (B Subunit)
girvfflknkygløkgmygenlelerllynsvoksyadrlfsyektkmyhdtplifwee
dkekcaeaekafiægokilldygksifwinendeinindpwssgimtvatrkvygevdds
erwhkvligklæddyeklleesskesteankklisdlvdrledaatkfirkgevetr
vkdlraryggtydprodteakkkvelleasletfldsieselvocledodiywkeodvkdl
artoeleeodøeakreeaaedlæsinerikksktmidrakwhienaedsitwhtsgiemk
dmkarlkilfeditsvlpeideietclsleelplittrelltrsylkfricsetlikmts
vfennivygevevolonigfklogisgargkkoddfanleedvalokkrireltonfeig
gfnfmkedfkaaakdlyirstaeokmnfdvpcmelfrryheevnkellelmyncadsyrd
akkklcslridekellokeikkeefyokkoorhadrsrhttyoklriaeelalelkkki

CPT_0013 18509 21106
pmp_J-Polymorphic Outer Membrane Protein
RDPKLAFFIYLLYWKESPLREKKVMKIPLRFLLISLVFTLSMSNLLGAATTEELSASNS
FDGTTSTTSFSSKTSSATDGTNYVFKDSVVIENVPKTGETQSTSCFKNDAAAGDLNFLGG
GFEFFFSNIDATTASGAAIGSEAANKTVTLSGFSALSFLKSPASTVTNGLGAINVKGNLS
LYDNDKVLIODNFSTGOGAINCAGSLKIANNKSLSFIONSSSTRGGAIHTKNLTLSSGG
FTLFQGNTAPTAAGKGGAIAIADSGTLSISGDSGDIIFEGNTIGATGTVSHSAIDLGTSA
KITALRAAQGHTIYFYDPITVTGSTSVADALNINSPDTGDNKEYTGTIVFSGEKLTEAEA
KEDKNRTSKLLQNVAFKNGTVVLKGDVVLSANGFSQDANSKLIMDLGTSLVANTESIELT
NLEINIDSLRNGKKIKLSAATAQKDIRIDRPVVLAISDESFYQNGFINEDHSYDGILELD
AGKDIVISADSRSIDAVQSPYGYQKWTINMSTDDKKATVSHAKOSFNFTAEQEAPLVM
LLWGSFIDVRSFQNFIELGTBCAPYEKRFWVAGISNVLHRSGRENQRKFRHVSGGAVGA
SYRMFGGDTLSIGFAQLFARDKDYFMTNTFAKTYAGSLRLQHDASLYSVVSILLGEGGLR
ELLPVYSKTLPCSFYGOLSVGHTDHRMKTESLPPPPTLSTDHTSWGGYWAGGLGTRV
AVENTSGRGFFQEYTPFVKVQAVYARQDSFVELGAISRDFSDSHLVNLAIPLGIKLEKRF
AEQYYHVVAMYSPDVCRSNPKCTTTLLSNQGSWKTKGSNLARQAGIVQASGFRSLGAAAE
LFGNFGFEWRGSSRSYNVDAGSKIKF

21922 21365 CPn 0014 CPT_UU14 21365 21922
pmp_3-Polymorphic Outer Membrane Protein
IONOSIYFTMKSSFPRFVFSTFAIFPLSMIATETVLDSSASFDGNKNGNFSVRESQEDAG
TTYLFKGNVTLENIPGTGTAITKSCFNNTKGDLTFTGNGNSLLFQTVDAGTVAGAAVNSS
VVDKSTTFIGFSSLSFIASPGSSITTGKGAVSCSTGSLSLTKMSVCSSAKTFQRIMAVLS POKLFH

CPn_0015 21935 24174
pn_3-PMP_3 (frame-shift with 0014)
LEFDKNYSLLFSKNFSTDNGGAITAKTLSLTGTTMSALFSENTSSKKGGAIOTSDALTIT
GNQGEVSFSDNTSSDSGAAIFTEASVTISNNAKVSFIDNKVTGASSSTTGDMSGGAICAY
KTSTDTKVTLTGNGMLLFSNNTSTTBAGGAIVYKKLELASGGLTLFSRNSVNGGTAPKGGA
IAIEDSGELSLSADSGDIVFLGNTVTSTTPGTNRSSIDLGTSANMTALRSAAGRAIVFYD
PITTGSSTTVTDVLKVNETPADSALQYTGNIIFTGEKLSETEAADSKNLTSKLLOPVTD
GGTLSLKHGVTLQTQAFTQQADSRLEMDVGTTLEPADTSTINNLVINISSIDGAKKAKIE
TKATSKNLTLSGTITLLDPTGTFYENHSLRNPGSVDILELKASGTVTSTAVTPDPINGEK
FHYGVQGTMGFIVWGTGASTTATFMWTKTGYIPNPERIGSLVPNSLWAMFIDISSLHYM
ETANEGLQGDRAFWCAGLSNFFHKDSTKTRRGFRHLSCGYVIGGNLHTCSDKILSAAFCQ
LFGRDRDVFVAKNCGTVYGGTLYYGHDFTVISLPCKRPCSLSYVPTEIPVLFSGNLSYT
HTDNDLKTKYTTYPTVKGSWGNDSFALEFGGRAPICLDESALFECYMPFMKLOFVVAHQE
GFKEGGTEAREFGSSRLVNLALPIGIRFDKESDCQDATVNLTLGYTVDLVRSNPDCTTTL
RISGDSWKTFGTNLARQALVLRAGNHFCFNSNFEAFSQFSFELRGSSRNYNVDLGAKYQF

CPn_0016 24383 25188

pmp_4-Polymorphic Outer Membrane Protein

RSDFALKROCHMRSSFSLLLISSSLAFFLLMSVSADAADLTLGSRDSYNGDTSTTEFTPK

AATSDASGTTYILDGDVSISQAGKQTSLTTSCFSNTAGALTFLGNSFSLHFDNIISSTVA

GVVVSNTAASGITKFSGFSTLRMLAAPRTTTKGAIKITDCLVFESIGNLDLINENASSENG

GAINTKILSLTGSTRFVAFLANSSSOQCGAIYASGDSVISENAGLISFGNNSATTSGGAI

CAECNLVISNNQNIFFDGCKATTIKGAILDCNKAGANPDPILTLSGNESLHFLNNTAGNSG

GAIYTKKLVLSSGRGGVLF:NNYAANATPKGGAIAILDSGEIISJADLGNIIFEGNTTST

TGGFAGSYTRNAIDLASNAKFLNLBATRGJKV/TYPOPITSSKATDKLSLIKRADAGSGNTYE

GYTVPSGEKLSEEELKKPDNLKSTFTDAVELAAGALVLKDGVTVVANTITQVEGSKVVMD

JGGFBGSTGABGVPLLNGIAINLDGLDGTHKAILKAGAKBVALAGGIMMUDAGGNTYEHH
MICCOLONGERLISEGRGVENTITELUDEDELJERFNIKKOVALAGGIMMUDAGGNTYEHH
MICCOLONGERLISEGRGVENTITELUDELDELJERFNIKKOVALAGGIMMUDAGGNTYEHH-CPn_0016 24383 25188 NEGOCOVERE EREGACCIMITITE ELIDTE FEBITINHY GYOCINWANG EGIGE ONCKNIKK CYEN

CPD_0017

CPD_00

73

CPn_0018 275.13 29003
pmp_5-Polymorphic Outer Membrane Protein
EYMMKTSVSMLLALLCSGASSIVLHAATTPLNPEDGFIGEGNTNTFSPKSTTDAAGTTYS
LTGEVLYIDPGKGGSITGTCFVETAGDLTFLONGNTLKFLSVDAGANIAVAHVQGSNNLS
FTDFLSLVITESPKSAVTTGKGSLVSLGAVOLODINTLVLTSNASVEDGGYIKGNSCLIG
KINSAIFFONTSSKKGGAISTTOCLTIENNLGTLKFNENKAVTSGGALDLGAASTFTAN
HELLEGGRETOSHAANE VALHE SERSTEINFELLDSBETMGYVAASTETTGICSLOFTE
GGDIVEGKOVTTTAPNATTKRNVIHLESTAKWTGLAASQGNAIYFYDPITTNDTGASDN
LRINEVSANQKLSGSIVFSGERLSTAEAIAENLTSRINQFVTLVEGSLVLKQGVTLITQG 27513 FSQEPESTLLLDLGTSL

CPn_0019 29007 30356

pmp_5-PMP_5 (frame-shift with 0018)
ASTEDIVITNLSINADTIYGKNPINIVASAANKNITLITGTLALVNADGAFYENHTLQDSQ
DYSFVKLSPGAGGTIITQDASQKPLEVVAPSRPHYGYQGHWNVQVIPGTGTQPSQANLEWV
RTGYLPNPERGGSLVPNSLWGSFVDQRAIQEIMVNSSQILCQERGVWGAGIANFLHRDKI
NEHGYRHSGVGYLVGVGTHAFSDATINAAFCQLFSRDKDYVVSKNHGTSYSGVVFLEDTL
EFRSPQGFYTDSSEACCNQVVTIDMQLSYSHRNNDMKTKYTTYPEAQGSWANDVFGLEG
GATTYYYPNSTFLFDYYSPFLRLQCTYAHQEDFKETGGEVRHFTSGDLFNLAVPIGVKFE
RFSDCKRGSYELTLAYVPDVIRKDPKSTATLASGATWSTHGNNLSRQGLQLRLGNHCLIN
PGIEVFSHGAIELRGSSRNYNINLGGKYRF

CPn_0020 32717 30603

Predicted OMP [leader (14) peptide: outer membrane]

KLWSMPNLRLMKRCFLFLASFVLMGSSADALTHOEAVKKNSYSLSFKSVSGIVTIEDGV

KLWSMPNLRLMKRCFLFLASFVLMGSSADALTHOEAVKKNSYSLSFKSVSGIVTIEDGV

KINIHNNLRIQANKVYVENTVQQSLKLVAHGNNWNYRARTLVCDYLEYYEDTDSCLLTNG

RFAMYPWFLGGSMITLTPETTVIRKGYISTSEGPKKDLCLSGDYLEYSSDSLLSIGKTTL

RVCRIPILFLPPFSIMPMEIPKPPINFRGGTGGFLGSYLGMSYSPISRKHFSSTFFLDSF

FKHGYGMGFNLHCSGKOVPENVFNMKSYYAHALAIDMAEAHDRYRLHGDFCFTHKHVNFS

GEYHLSDSWETVADIFPNNFHLKNTGPTRVDCTWNDNYFEGVLTSSVKVNSFONANQELP

YLTLRQYPISIYNTGVYLENIVECGYLNFAFSDHIVGENFSSLRLAARPKLHKTVPLPJ

TLSTLGSSLIYYSDVPEISSRHSGLSAKLQLDYRFLHKKSYTQRHHILEPFVTFITETR

PLAKNEDHYIFSIQDAFHSLNLKAGIDTSVLSKTNPRFPRIHAKLWTTHILSNTESKPT

PEKTACELSLPFGKKNTVSLDAEWWKKHCWDHMINTWEMIGDNVAMTLESLHRSKYST

IKCDRENFILDVSRPIDQLLDSPLSDHRNLILGKLFVRPHPCWNYRLSLRYGWHRQDTPN

YLEYQMILGTKIFEHWQLYGVYERREADSRFFFFLKLDKPKKPPF

CPM 0021 34470 32707
Predicted OMP [leader (19) peptide]
CSREPYPNIEILARGVEHRSMGLFHLTLFGLLLCSLPISLVAKFPESVGHKILYISTOST
QOALATYLEALDAYGDHDFFVLRKIGEDYLKQSIHSSDPOTRKSTIIGAGLAGSSEALDV
LSQAMETADPLOQLLVLSAVSGHLGKTSDDLLFKALASPYPVIRLEAAYRLANLKNTKVI
DHLHSFIHKLPEEIQCLSAAIFLRLETEESDAYIRDLLAAKKSAIRSATALQIGEYQOKF
FLEPTLRNLLTSASPODQEAILYALGKLKDGOGSYYNIKKQLOKFDVDVTLAAAQALIALGE
EDALPVIKKQALEERPRALYALRHLPSEIGIPIALPIFLKTKNSEAKLNVALALLELGC
DTPELLEYITERLVQPHYNETLALSFSKGRTLONKRVNIIVPODPOERERLLSTTRGLE
EQILTFLERLPKEAYLDFIYKLLASQKTQLATTAISFLSHTSNGEALDLLFQAKLGEP
IIRAYADLAINLTKDPEKKRSLHDYAKKLIQETLLFVDTENQRPHPSMPYLRYQVYPES
RTKYALDILETLATSKSSEDIRLLIQLMTEGDAKNFFVLAGLLIKIVE

CPn 0022 35042 34395

maf
TIEQVISNCCNVSNTRSFYSMSLPLVLGSSSPRRKFILEKFRVPFTVIPSNFDESKU
GDPLAYTQELAAQKAYAVSELHSPCDCIILTGDTIVSYDGRIFTKPQDKADAIQMLK
NQTHDVVTSIAVLHKGKLLTGSETSQISLTMIPDHRIESYIDTVGTLNNCGAYDVCH
ILKKVHGCVYNVQGLPIQTLKYLLEELNIDLWDYSI

CPh_0023 36657 35014 yj5K/alr-ABC Transporter Protein ATPase yjjk/aif-ag. Transporter frocein aifase
engakliyskohtymisamsividkigkslotrilfddvsvvfnpgncygligpngagks
tlliingmieptrgsislpkkvgilronidsfhdttvldcvimchtrlwealdyrdnly
loeftdaigmelgeieeiigeemgyradseaeelligigipnemfokmamipydlofrv
Llcöalfghepallldeptnhldlysinwlghflkdyegtvivvshdrhfuhttthiad
ldvetiiiypgnyddmvemktasreqekadikskekkisolkefvakfgagspasovosr LREIKKLOPOELKKSNIQRPYIRFPLSDKSSGKVVLSLEAITKDYGDHQVIHPFSLEIYQ GDKLGIIGNGLGKTTLMKLLGAVEAPSSGSIKLGHQAICSYFPONHSDVLADCOEPTLF EWLRNKTGINDGEIRSVLGKWLFGGDDAFKGIQALSGGETARLLMAGKMYENHVLILD EANNHLDLESVSALSWAINDYKGTAIFVSHDRGLIQDCATKLLIFDKDKJFFFDGTWVDY

37605 36661 xerC-Integrase/recombinase Xett-Integrase/recombinase

REWMIASIYSFLDYLMWKSASPHTLRNYCLDLNGLKIFLEERGNLAPSSPLOLATEKRK
VSELPFSLFTKEHVRMYIAKLIENGKAKRTIKRCLSSIKSFAHYCVLOKILLENPAETIH
GPRLPFSLFTKEHVRMYIAKLIENGKAKRTIKRCLSSIKSFAHYCVLOKILLENPAETIH
LSTHLIRIRICKGKKERIIEVTSNAIOWIGIYLNHPORKELEKDPOAIFLNREGRRISTRS
LDRSFOEYLRRSGLSGHITPHTIRHTIATHWLESGMDLKTIOALLGHSSLETTTVYTOVS VKLKKOTHQEAHPHA

CPn 0025 38610 37684 CPT_0025

J8610

J684
elaC/atsA-Sulphohydrolase/Glycosulfatase

LLMSSRELIILGCSSQQPTRTRNQGAYLFRWNGEGLLFDPGEGTQRQFIFANIAPTTVNR

IFVSHEHGDHCLGLGSMLMRLNLDKVSHPIHCYYPASGKKYFDRLRYGTIYHETIQVVEH

PISEEGIVEDFGSFRIEAQRLQHQVDTLGWRITEPDTIKFLPKELESRGIRGLIIQDLIR

DOEISIYGSTVVLSDVSYVRKGDSIAIIADTLPCQAAIDLAKNSCMMLCESTYLEQHRHL

AESHFHMTAKQAATLAKRAATQKLILTHFSARYLNLDDYYKEASAVFPNVSVAQEYRSYP

CPn_0026 39637 38762 CPU_0026 - 39637 38762 CTT45 Sypotherical Dede-En CNEAMBILLEGLEBOYTSYFIKEOPEKOAAPOKSTYDTCHTAYLELECVLV77VLVGAML CMETOWELEGLEBOTALLVLLSTEPPOLENWESYKETKETAEKOASESQFYKSASRKGS CQLEHABDHERKNETRYOLEKGVNYVTNKEKRSPEOPHLSDEBINGPROSKROSELESODE CDELLBPANGKABITATTKESKETSTTESSKKKKKKKISCHBTTCSTHKRSAPKPMVPSK KRKEVLJKKTVPLIFEDLEHQGGGNESSDESSPPEVORKATLEWPCKOPTDP

39778 CPn 0027 42252 Ton Lon ATP dependent Professe

PSIRTIVDSTTNSDSPILDI

CKLLDEGEEEGEDQSTERLLPSELFILPLNKRPFF
PCMAAPILIESGPYYEVLK.

OKYIGLVLTKKENADILKVSFNQLHKTGVAARILR
IMPIECGSAQVLLSIEERIRTTETIKDKYLKARVSYHADNKELTEELKAYSISIVSVIKD
LLKLNPLFKEELOIFLCHGDFTEFCKLADFSVALTTATREELQEVLETTNMHORIDKALI
LLKKELDLSRLQSINQKIFATITKSGKEFFLKECLKKIKKELGLEKEDRADIDIEKFSE
LRKRHVDPYAMEVIQDEIEKCOTLETSSAEYTVCRNYLDWLTIIPWGIQSKEYHDLKKAE
IVILKDHYGLDEIKOR ILBLISVGKLSKGIKGSIICLVCPPCVGKTSIGRSIAKVLHRKF
PRFSVGGMRDEAEIKGHRFTYIGAMPGKMVOALKOSQANNPVHIDEVDKIGASYHADPA
AALLELDSBENKGELF, FLISBOLGHSTELGTHIDHINTAREAGVETLIFBIELGTYLLE
BETLIABETTIVTARPFIRITALSEVIEFITSAEFTMINNYAREAGVETLIFBIELGTYLLE
BETLIABETTIVTARPFIRITALSEVIEFITSAEFTMINNYAREAGVETLIFBIELGTWILE
BETLIABETTIVTGAEFTMISSCIAWTYLHSALHRYAPGYTFFPKSQVHIHIPE
GATPKDGPSAGITMYTSLLSLLLETPVVNNLGMTGEITLTGRVLGVGGIREKLIAARRSR
LNILIFPEDNRRDYEELPAYLKTGLKIHFVSHYDDVLKVAFPKLK

CPn_0028 43328 42543
No robust homolog present in Genebank/EMBL as of 11/7/98
RMFLOFFHPIVYSOSLSFLPYLGKSSGIIEKCSNIVEHYLHLGGDTSVIITGVSGATFL
SYDHALPISKSEKIIKILSYILILELILALFIKIVLRIILFFFYRGLILDVKKEDLKKTL
TPDQENLSLPYPSPTTLKKIHALHILVRSGKTYNELIQEGFSFTKITDLGGAFSPKQDIG
FSYNSLLPNFYFHSLVSVPNISGEERALNYHKEQQEEMAVKLKTMQACSFVFRSLHLPSM
QTKDKKAGFGLLTFFPWKIYPL

CPn_0029 43390 LOEITWSRSVMELKIKFYVYLNSERNKTKP

CPn_030 43840 44529
gcp-0-Sialoglycoprotein Endopeptidase
LKGVCWYSLFFY INNRHYFYKYVIIDTSGYYPFLACVDNQQVLEHWSLPVGPDLGIVLE
FLYKSWILSFGQVAULGPGNFSATRIGISFAQGLAMAKNVPLLGYSSLEGYLLSKDEKK
AVMLPLGKRGGVLTLSSEIPEEGLNEKRRGVGPGALLSYEEASDYCVAHGYYHVISPNPQ ASSFSDKITVEEVAPSVEQIRRHVISQFMFVEYDKQLSPDYRSYSCIF

44708 44884 CFI_0031 rs21-s21 Ribosomal Protein CMPSVKVRVGEPVDRALRILKKKIDKEGILKAAKSHRFYDKPSVKKRAKSKAAAKYRSR

46098 CPn_0032 44881 CPH_0032 44881 46098

dnaJ-Heat Shock Protein J

SLICHNVFVGSVSGMDYYSILGISKTASAEEIKKAYRKLAVKYHPDKNPGDAAAEKRFKE
VSEAYEVLSDPOKRDSYDRFGKDGPFAGAGGFGGAGGMGNMEDALRFFMGAFGGEFGGGS
FFDGLFGGLGEAFGMRSDPAGARGGASKKVHINLTFEEAAHGVEKELVVSGYKSCETCSG
GAVNPGGIKSCERCKGSGQVVQSRGFFSMASTCPECGGEGRIITDPCSSCRGGGRVKDK
RSVHVHIPAGVDSGMRLKMEGYGDAGQNGAPSGDLYVFIDVESHPVFERRGDDLILELPI
GFVDAALGMKKEIFTLLKTEGSCRLTVPEGIQSCTILKVRNQGFPNVHCKGRGDLLVRIS
HERDONI EBERNYEI LEMPA CTEVA FINDPKKBSET DKYKKEFESFTV VETPONLSEEQKELLRTFASTEKAENFPKKRSFLDKIKGFFSDFTV

CPn 0033 46129 48171 pdhA&B/odbA&odbB-(pyruvate) Oxoisovalerate Dehydrogenase Alpha & Beta Fusion ERSMGYVONOVISSIRDVLKLVWELRFAEHKMLLLSROSGSGGTFOLSCAGHELAGVLAG KSLIPGKDWSFPYYRDQGFPIGLGCDLSEIFASFLARTTPNHSSARMMPYHYSHKKLRIC CQSSVVGTQFLQAAGRAWAVKHSSADEVVYVSGGDGATSQGEFHEMLNFVALHQLPLITV CQSSVVGTQFLQAAGRAWAVKHSSADEVVYVSGGGGATSGGEFHEMLNFVALHGLPLITV IQNNHWAISVPFEDQCGADLASLGRCHQGLAVYEVDGGNYTSLTETFSHAVDQARQHSVP ALILIDVVRLSSHSNSDNQEKYRSALDLKLSMDKDPLILLEKEAINVFGLSPFEIEEIKA EAQEEVRKSCEIAEALPFPSKGSTSHEVFSPYTETLIDVENSESAQNLRNSEPKVMRDAI SEALVEEMTRDSGVIVGGEDVAGDKGGVFGVTRNLTEKFGPQRCFNSPLAEATIIGTAIG MALDGIHKPVVEIOFADYIWFGINQLFSEASSIYYRSAGEWEVPLVIRAPSGGVIQGFYHSGSGSVEGFLAHCPGIKVAYPSNAADAKALLKAAIRDPNPVVFLEHKALYQRRIFSACPVFSHDVVLPFGKAAIVHPGKDLTIVSKGMPLVLSLEVAQELASRGISIEVIDLRTMVPCDFA TVLKSLEKTGRLLVIHEASEFCGFGSELVATMSEQGYAYLDAPIRRLGGLHAPVPYSKVL ENEVLPHKESILQAAKSLAEF

49496 48210 CPT_0034 49496 48210
CT345 hypothetical protein
vnFLLPTTCRGILMAEISTPSLPDSSIVSQKTPPVPDPDSSPDHIPTIPTQAPFKPQRKK
ETPSSIVNAIAFAILAFLSCLGGVFAICLGCSLEITMPLFILTAVFIAFTLLYFIHYLEK
PKIFPEDLPTPPPSYTLRAPTLTPEIPAPAPGIPLPPTLPKVDRTKLTCNPDIHYPSTYDF
KACFSLLKQLFSLDPETRPEDRKYSNKLASILLRSKEKSGFRFHCFKGHFSHDKILNKKS
GAVVISSHSSMDFSTTLGRAFAVTTCLQRSCWEKIKNNIPTPEKHLPIGSCVSGPMDVEE
GAQLYTSHLVINPPTLETLIKEKMRRAITLKDFSMKEAFTHUVLAYLQFDICTEHNLE
SVOLEVFGLNNLSADQEEFTTWESCCHLALLESVRILLASKEEYALSNVSVNSISQVPLQ TACRALFLN

CPn 0035 51146 49569 CTT339 hypothetical protein
ARTTLEEDAGSSLKPLPKTFPCATALYITHRRERKSEHQMWNRCQVFSSFFFRYPISSWL IRLRASCECFQQRHPIFLCGLYWLAGITSRCHPECSALILIFLCMFLPRNPKQMLPLASA WIISLMLTPAPFLHDGPISGTFVIHHAGGQTTYYGEALCIQTPCGKRAHHLSCQILSESR LELKKVYELEGTLHHTSQIVFKSNACYKEIPRSRFYIMKEKCRESSCHFLNHRFPSSEVG PFASSLLLCTPLPONLRDLFROKGLSHLFAISGWHFSLCATTLWMLCALLPLKIKKILSF IVLTSLACIFPMSLSVWRSWISVTLLCFSWCFSGSCSGLNRLGAGFILCSIFFSPFSPTF VLSFLATIOILLFFPKTFSFLYTPWTQFLSPFWLYPIKYLAMTLAISLSQCFIVTPMQ YFGSLPLEGLLYNLIVPFTTLPIIVFLIATIILPCCSPITEALIQGFLSHPWLHNPNILK TLSFAPVPPWMLTLASLILFFIGILRTNVSPYASISATSYRFIETL

CPI_0036
SINGS TO STANDARD TO HIGGLET FELLHLUKGYYPOFLOWLENYLHGEEVYNE

51702 AGENTALIST PROSPHOGATE FOR PROSECTION HOR KLIGOTECHRALILLISTROWRTTY (VERCIMMERTRTYLESEKDTYDD) FEELYATTIVKN AGENVERDEN VERLETGEN TYMETYATKTINAKSIMSILMEVARYXSELLYTIRDREA

HRILLOK LODAFSSGFGEL

CEn_0038 CPI_003H
TOTAL Hypothetical protein
MDTOSSTGMEENRIAGTSIVSGMALGKVFFLGTSPLHVRELTLPQEEVEHEIHRYYKALN
RGKSDIVALEQEVTQQQLQEVGSILQAHLEIMKDPLLTEEVVNTIRKDRKNAEYVFSSV

MGKIEESLTAVRGMEGVVDRVQDIHDISNRVIGHLCCQHKSSLGESDQNLIIFSEELTPS EVASANGAYIRGFVGLVGAATGHTAIVSRAKSIPYLANISEELWNIAKRYNGKLVLIDGY EVANANIATI KAYANIA ATALAHAN YAMITSIPPLA PETILIPI NEBIHA E CECEPENLEMENENTA IN MERIHANDAN TALAHAN MERIHANDI MENANGKAN NSCLEITEALLQAKTCSEVEELLNRNNKITS

CPn_0039 54256 53963
CT339 hypothetical protein
ISMOSGYAKKKEAKTHEQOFLEMEASLLEKRYEGQAGNGLVSVVINGKCDLISVKVQPT
CLDPEDPEVIEDLFRAAFKLAKEQMDQEMSLMRSTMPF

CPn_0040 55673 54318
dnaX-DNA PO1 III Gamma and Tau
AFYTHSLGYTMTLQPYQASSRKYRRQIFREILGQSSVVAVLKNALVFNRAAHAYLFSGIR
GTGKTTLARILAKALNCVHLSEDGEPCNQCFSCKEIASGSSLDVLEIDGASHRGIEDIRQ
INETVLFTPVKAKFKIYIIDEVHMLTKEAFNALLKTLEEPPQHVKFFFATTEIHKIPGTI
LSRCQKMHLQRIPEKTILEKLSLMAQDDHIEASGEALAPIARAAQGSLRDAESLYDVYIS
LFPKSLSPDTVAQALGFASQDSLRTLDNAILQRDYATALGIVTDFLNSGVAPVTFLHDLT
LFYRNLLLTNSTTSKFSSQYKTEQLLEIIDFLGESAKHLQNTIFEQTFLETVIIHIIRIY
QRPVLSELISSIKSRQFEGLRNIKEPTLTQQVSAPQPQPTYKEQSFLEKKNQPAAEGKII
SVEVKSSASIKSAAVDTLLQFAVVEFSGILRQ

CPn_0041 55888 57342

CPn_0041 55888 57342
No robust homolog present in Genebank/EMBL as of 11/7/98
CKYLYHHSYPPPQHSVGSISSRYKLRVLAITFLVLGVLLLISGALFLTLGIPGLTAGVSF
GLGIGLSALGGVLVVSGLLCLLVKREVSKVCPEEIPAVOPEETPEGVPVTPFEKPALDEA
QKEQKTQKILDQLPQELDQLDRYIQEAFACLGPLKDLKYEDQGFLQDVKEEFQVFDFVQK
DMIAEFVELQQILCQEGRLLEFVINQTRYIGRDLFKREDSLYKLMEMLGYLPSGDVRGER
LKKSAREVVDRFMRTTCNIRKIAMTFDRHVYSVAKTAFEKAFGALETCVVESKRESYREA
FCEYEKAKLLGDEEKSAHAEORFGDIKGNWEDVKDAFFWAKEGKIEIDAGIGNSCWAE
RYEEHRITRARWYKVAEHQLFNATMRVKDSLREHNEARVAFEKERSKENQRQVQKKKEKR
LRÖEKELHDQELPRAQERLRELQALYPEIAVSVVEARREVASDLEKAHESIDKHYQSCVR
EÖEEY

ENDEKELHDUELPRAGERIKELGALIFELAVSYLLINGER

DOTLY

ETE

CPT-0042 57346 58182

NO-ECOBUST homolog present in Genebank/EMBL as of 11/7/98

EEEEKQEAEFRENGTKIRSMEEVSEYLQOVENQLESCSKRLITKMETFALGVRLEAKEEIE

STENSOVNRFEVLCRDIEDMLSRVEEIERMLMABLPILLPIKEAITKASTVQINSCKEKL

TKVEPYFKESPAYLTSEERLQSLNQTLQRAYKESQKVSGLESEVRACREQLKDQVRQFET

QGVSLIKEEILFYTSTFRIKFSYHSFRIHVPCMRLYEEYYDDIDLERTRARWMAMSERYR

DARGAEDGULVEGTUEFAGAIRFETVULVREFREKSKKKH DAFQAFQEMLKEGLVEEAQALRETEYWLYREERKSKKKH

CHIDO43 58432 60372

NO CODUST homolog present in Genebank/EMBL as of 11/7/98

HHRFIMOVPLSPOLPPPPPBHSVGASFCLSKFRVLAITFLVLGVILLISGALFLTLGISG

VSLGVGLGLSALGSVLVISGFLLLLERREVSGVGLEGIPTGIPVGPSAEPSSEIGN

KNOILDQLPQELDQLDTDIQHVLSCLGKLKDLKCKDRGLLKDAKERLQVFDFVWKDMME

FVELQOVMDQESRYLEGIIHEVQSIAHKLFVDDVNIRSHLGESGGYLPSEDVRGELLKRF

AKEWARFMKVTRDIRKIAMAFNKNAYGAAKNAFDKAFGSLETCLYKSLTKSYRDTFCDY

KRAKILPDENNSARAEQAFREVKDHWEDLKETVFWKEDGRIDIEVLTAVGGWPDRYPEL

LILETKRKDKVMSHQLWEATHMVKEAEVTYSVARVAFEKDGSQONGKFGEKTKERLRCLK

DLRDDECHRAQERLEKLTALYPEVSVSVVETEREBKFNLEKAYGNLEERYGSVVQDQEDY

WTEÖRNREAEFRAKGTKVRSMEEVAEHLQILENLLEDCYKRLSKAETFALGVEREATEEI

EYTLSDAANRLKVLCEDIEDTLPRVEEIEMMLRMAERPLHPIKQAFTKAFVQYNRCKER

LAKVEPYYKESPAYVNSEERLQSLDQASQCIQRVPKGFKFRNGSMYI

CPd_0044 60278 60778

NOAKOUST homolog present in Genebank/EMBL as of 11/7/98

IAKSDCRWWIRLHSAYKESQKVSSLETEACTYREYLREQVQOFETQGVSLIKEELLFLSS

TLKSKLSYDPLIANIPCMKFYYQYYDDIDKARAQSRWLEKSERYRNAKRRFQEIVKKGLF

KEAKPLKKEEYRLLQEERSNKEKRLIYNKMAVARQRVQEFESMEIPE

60961 62790

CPn_0045 60961 62790
CT345 hypothetical protein
CKYTYHPPOLPPDHSVGATSWOPKLRILTITFLVLGVLLLISGALFLTLGVPGLAAGLSF
GLGIGLSALGGVLVVSGLLFFLIRRGVSKVRPEEIPVTPSHEAGKILGQLPGELDOLDTS
IQEVVSCLGKLKDLKYEDQGLLTEVQEKLRVFFFVRKDMVTEFLELQQVVAQEGOYLDYL
INQVGSISHKLFVPDVIGAHLAELGGYLPSGDVRVERLKRSARQVVDRFMRVTYDTRKV
AMAFDENACGVAKNAFDKAFGALBECVYKSLTESYPEAFYEYSKAKILRNEDVFALQDKN
KSARAEQRFREVKDRWEDLKETVFMVKENGCIDLEVLTAVGGMPDRGPEHLIPSKRRNKV
MSHKLWEATMRNKGAGGTYSVARVAFEKDGSRKNQKKFQEKTKEWLRCLKDLHDGECHRA
RERLAELEALYPEVSVSVVETERETKFKLETAYGNLEERYQSVVRXQEDYWKEEENKEAE
FREKGTKVRSPEEVVEYLQILENLSEDCSKQLTIAET/VLGVELEATAEFFYTILSDAAN
RLKVLCEDIEDILPRVEEIEIMLRIAELPFLPIKQAFTKAFLQYNSCKDKLAKVEPYCQE
SVDYKSGFRV SVDYKSGFRV

CPn_0046 62775 63263
NO robust homolog present in Genebank/EMBL as of 11/7/98
ERFOSLM)DLONVYQECOKATOLESEVSAROHLERGITEFETOSLTVIKEELLEVSSTL
KSKLGYOPLIADIPCMKFYEEYYDOIDKARVOSRULEYSERYKKAKGFQEMLKEGLFKE
DQALKKAEYRLLREKRMNKEKLLICNKIEAAQQRVQEFGPSDS

CPT_0047

No iobust homolog present in Genebank/EMBL is of 11/7/98
EMERICANTIVEFREVLEGILITMYHFOKIRMTLITT/GF7/LNKYLRKDVELWFVYGSCPES

63687 CPOPCEDAIMDILKIANIES
ASCCVKGELLGKRCIEKITKOTPILEKYORIDDRD
AKILKQLRAQLLSVHTLES
AIFVVLLILLWGYGALKALCEMLKSPORFHLYIA
ILITLSLLWCROTEEIFCAYMVATESYPPILPETAVLLGYFLGLEIAGFSCTFLALLYTLGS
DLWNNSWFLSINLLCSWRILVSLHRVSRLSSVEWAGMKLGOVANGSLLURFIFITNTISRE
ALYADGIESFVYSLITAISVVALIEVFEASFGASTNFSLLTYLSPENALLKRLFKEAROT
YOHSVLVGSLAEAAAQAIGABSLYCLVAAHYHDYCKLINNGFFSENGKILQOSGHSLSPL
CGAMMINHNIPEGVALARQAGLPESFIGVIEEHGGTSVIRSAYYSHWVENPSTGSFDEEL
FRYSGNKPSSKETTIIMIADSFEAASRSLKNASLPDLQRLIDQIIQGKLQDGGFSCSPIT
LOBLALLIBKONGOT WALHIGPMMYDELLYG AMDUSTEDSIGGT

No robust homolog present in Genebank/EMBL as of 11/7/98
LKEKRRNIVYLLVIYQEIFMLTMLHQPYYDKILTCMTIYIPGHTHKDSNKLFQKKSRAIW
VDEKPFSLDCFSNVFLIFVSLVPIAGLYRAYQIKKSLDRTTVQIGYSPSLSCEQKECVEA FVNGYGLICISILGGLGILVPILILVVLSLLLLGILMLFSLSTYESIKNYISKHICWKSN

CPn_0050 66849 66499
No robust homolog present in Genebank/EMBL as of 11/7/98
VSWFPILGIFLAMRYAKHOTNWNDENVKANLGYLPSTNCKNALCRMSSRLTSSIKTAGIL
GGGGILLPIFILLLAILLISVLBQLIMLPFRLCCFALRGSVSSDTVTNLLLLANTILA

66797 67111 No robust homolog present in Genebank/EMBL as of 11/7/98 CFAYLIARNIPRMGNHETYIAPGVLPSSHAQDVSRSTVYPSRSFIMRRMLMGWNFNRVPS KSSEQLMDGHRIPLIFFGKHPTISILNVNRFSWLSIFYNGERGF

CPn_0052 /68008 67304
hemC-Porphobilingen Deaminase
whLsvCySDPCLSDFC0cKrPLRIASRNSNLAKAQVHECISLLRSWYPKLWFQLSTTETT
GDREKK IPLHLVENSYFFTDGVDALVHKGVCDLAIHSAKDLPETPSLPVVAITRCLHPAD
LLVYADHYVHEPLPLSPRLGSSSLRRSAVLKQLFPQGQILDIRGTIEERLDQLHRGHYDA 68008 IVLAKAASLRLHLHAAYSIELPPPYHALQGSLAITAKDHAGKWKQLFTPIHCHSS

CPn 0053 69350 CPN_0053 / 69350 67986
sms-Sms Protein
IRMATKTKTOWTCNCCGATAPKWLGQCPGCHNWNSLVEEYVPQARSGTSSRSSTSAIALS
SIELENESRIFIDHAGWDRILGGGVVRGSLTLLGGDPGIGKSTLLLQTAERLASQKYKU
VCCGESVTOTSLRAKRINISSPLIYLPFETNLDNIKQQIATLEPDILI IDSTQIIFNPT
LNSAPGSVAQVREVTYELMQIAKSAQITTFIIGHVTKSGEIAGPRVLEHLVDTVLYFEGN
SHANYRNIKSVKNRFGPTNELLILISMHADGLKEVSNPSGLFLQEKTGPTTGSNIIPIIEG
SGALLIEJALVSSSPFANPVRKTAGFDPNRFSLLLIAVULEKRQVKLFTMDVFLSTTGGL
KIIEPAAGLGALLAVASSLYNRLLPNNSIVIGEVGLGGEIRHVAHLERRIKEGKLMGFEG
AILPEGDISSLPKEIRENFRLQGVKTIKDAIRLLL

0654 70089 69313

CPILJO34 69313
ric_Ribonuclease III
rLSFFPPIKIPNSKFKDGALLSMHPPIDITAIEAKLNFTFTOPKLLEIALTHPSYKNESA
QQEDSERLEFLGDAVLGLIVTEHLFLLFPSMDEGTLSTARASLVNAKACCRYTTMLGIG
FWLLIGKGEKIGSERGRUSAVANLFESILGAVYLDGGLSPARKLTVPLLPPREEILPLMS
GMPHNLLOOFTOKOFRVLPVYOSTAVTDAQGNVSYQIQVLVNQEVWGEGNASSKKEAEKI
MAQQALDTYGNKNQNTMDV

Cp_0055 70096 70590 Cf296 hypothetical protein CFWICYLIRIRWRSALHLOHLRHFHNHGSILFENLLTIKDCFLLETKLONFIAKASKTID DVRWRENIFRSMPEIYTVVRKRRLDFFAAELVHRPKLSLVRDLWVFFGEETLEGEECML LLSGDRAGSGIFFTGPYPSDLYELEKGTTGLLLAFSSVGIPVI

70917 72746

CPn_0056 70917 72746

mrsa-Phosphomannomiase

EFLKISLHRISLMKEVEQRIRSLYDAVTAENICRWLSNDCTQQDAKTILGWLDTDPAQLE

DLFGATLTFGTGGLRSLMGIGTNRINLFTIRRTTQGLVQVLRAHLPHPGDPMRVVVGCDT

RHNSIEFAQETAKVLAGNGCEVFLFQYEPELALVSFTVRYERAIGGVMITASHNPPNYNG

KVVMASGQQVLPPLDQEIVAACSAVNEILSVPSIDHPNIHLIGKEYEALYRDTLKQLQL

YPEANRISGRSLSISYSPLHGTGISLVPHVLKDMGFLSVHLVEKQAIGDGDFPTVQLPNP

EDPEALTLGIEQMLANDDDLFIATDPDADRVGVVCLEEDGPYRFNGNOMASLLADHILGA

WSKTRHLGEHDKLVKSLVTTEMLSAIAKHYHVDLINVGTGFKYIGEKIESWRNSTNKFVF

GAEESYGCLYGTHVEDKDAIIASALIAEAALQKLQGKTLCDALLSLYETYGYFANKTES

VVFSAKTDEQEIRKKLSHLEEISSANFFSGKYQVEKFENYKQGIGFNLLSKDSYALTLPK

TSMLCYYFSGGGRVIIRPSGTEPKIKFYFEMSTHYPERVTDKEIQKQREAESFQHLDDFI

FDFKEKFSNL

72913 73554 CPI_UO37 ... 3554
sodM-Superoxide Dismutase (Mn)
ILKRYVVMSFVPYSLPELPYDYDALEPVISSEIMILHHOKHHQIYINNLNAALKRLDAAE
TQQNLNELIALEPALRFNGCCHINHSLFWETLAPIDGCGGQPPKHELLSLIERFWGTMDN
FLKKLIEVAAGVQGSGWAWLGFCPAKOELVLQATANQDPLEPLTGKLPLLGVDVWEHAYY
LQYKNVPMDYLKAFFQIINWCHIENRFSEIISSK

CPH_UUSB /362/ /4562
accD-AcCoA Carboxylase/Transferase Beta
IRWLVRLFSYDKPKIKVQKIKADGFSGWLKCNHCHEMIHANELCONYNCCPKCSYHYRIT
AIERVKLLADKDSWRPLYTDLKSODPLEFIDTDTYANRLEKARKNTTESEGVIVGICTIG
LHPVALAWMDFNFMASSMGAVVGEKLTFLIEEAIETRLPVIIVSASGGARMQESVFSLMQ
MVKTSAALAKLHEAGLPYISVLTNPTSGGVTASFAALGDIIIAEPKALICFAGPRVVAQV
IGEDLFEGAOKSEFLLEHGMIDKIVERKELKTTLQTLLDYFLAQEYTGGKSKAPRDLSKR
/WETEIITDDSE LKEIFLLTDDSE

CPn_0059 74562 75050
dut_dutP Nucleot idohydrolase
tKHHTASCNONTIONATEMTYPGGEDPSGEEPETTPGAAGADDRAN LEEPTALLPYGRA
LIPMIPAE PEGYELDYRRSGLALKHGITYLIISPGTT LSDYRGE IRV ILINEGDSTPI
LEPKMP IAQVVLOPYVQATPVVKQEGLAETARGUGGGGITGAG

CPD_0050 75004 7552H
ptsh-PTS_IIA_Prote-in
PKLPEEYEVUTLEQAKMESYCOMOODFOLFSLLSPRLVMFLGKHSRDETLQULTDLVDA
AGLLEDKQAFFDALVRREN INSTS]GWZVATPHOKLESCONFFTATGTITTQSTLWDATDG
ALVRLVFLTGGPENAQAEYLKLLSTLTLSLRKESPROOLLQVNTTEEVMNVFVGM

75501 CET_UDGT /5501 / 6208
ptsN-PTS (IA Protein + HTI DNA-BINDIT Jin
RSHECICGDVKMDLKLDEVASLLDVSEHTVLQWLKEG. SMNNEYRFSREEIENWLL
HNOALH QOERGEDKEALKOLSLKYSLYKAIHRGGVLCDVVHSKEEALQYASKYIAQKFQ
LDESVLFEMLSHERICMSTGIGGGIALPHAKDFLINAYYDIVVPHIAEPIETYGALDGKP
VGILFFLFACODKSHLNLVNKIVHLGMSLNARSFFKNYPNKDQLLAYVKEWESQTH

CPD_0062 76251 77690

CTOSC TOP OF THE PROPERTY OF THE PROPERY

CPn_0063 78109 78267 No robust homolog present in Genebank/EMBL as of 11/7/98 PMYANCKHNCLCLYDFSRHRSPPGLPLTFTPPYSFTLGIFLGRCLSTSNIVLL

78340 78576 CPn 0064 No robust homolog present in Genebank/EMBL as of 11/7/98 LVMTKIQCSAQYYRSRPAERAQTPPQPFLARDRADFWERHPRFSACCRVLLLVAWVVLAL LFLFVMLLPLAAGSYLLAF

CPn 0065 78882 80651 CT288 hypothetical protein
YDYYKYNMFFKKNYMTDFPTHFKGPKLNPIKVNPNFFERNPKVARVLQITAVVLGIIALL sgivliigtplgapismilogcllasggalfvogtiatiloarnsykkavnokklseplm Erpelkaldyslolkevwolhhsvvkhlkklolnlsktorevlnoikiddbgpslgecaa ERPELKALDYSLÖLKEVMDLHHSVVKHLKKLDLNLSKTQREVLNQIKIDDEGPSLGECAA
MISENYDACLKMLAYREELLKEQTQYQETRFNQNLTHRNKVLLSILSRITDNISKAGGVF
SLKFSTLSSRMSRIHTTTTVILALSAVVSVMVVAALIPGGILALPILLAVAISAGVIVTG
LSYLVRQILSNTKRNRQDFYKDFVKNVDIELLNQTVTLQRFLFEMLKGVLKEEEEVSLEG
ODWYTQYITNAPIEKRLIEEIRVTYKEIDAQTKKMKTDLEFLENEVRSGRLSVASPSEDP
SETPIFTGGKEFAKLRRQTSQNISTIYGFDNENIDPEFSLPWMPKKEEEIDHSLEPVTKL
EPGSREELLLVEGVNPTLRELNMRIALLQQLSSVRKWRHPRGEHYGNVIYSDTELDRIQ
MLEGAFYNHLREAQEEITQSLGDLVDIQNRILGIIVEGDSDSRTEEEPQE

GPM_0066 80916 82655
No=robust homolog present in Genebank/EMBL as of 11/7/98
GYMANPTOSRPSPEISIEELELOELAGSSNTETISNTPPPSCAATAEEVSLFIEGGRR
NSEDEEGFLGSCEVYDVCITNGGDPEVRDHEVRVMYINGSGRTOHEGILDAMNICDLRG NSEDEEGFLGSCV7DWCITTNGGDFEVRDHEVRWYYINGSGRTOHEGILDAMMICDLRG
EBYRFIHNSGYGLGSCFLGIRNIPPRDNVISQAIQARWHEFIFAENANDDY LVLFSGN
GGLYLOVALDNSIYSHHILCVGIGSSYYIQGNYRVHNYRVTGDWTTLLDRRGATAVNTTT
LIMADSAEGLFLPSVRCPSYQWALRCGEOCLIMDNNQOVGFREODSSEIALWWNLNODH
STATRLIEW IDRGDSQAVLELNPOPSHCRDIALTALYATTRISSLLQBCLMISTYTAPEDY
FARVAIVTGYSIMTLRYFILLLTNRFGCRRHFRVLRLAALGLQSLGFLTVLLDHTNVTRR
WIRRPPLISVYFCTASFATGSFIYVDLTRWFFTSLRSRLQLFVORRLTGRGLFLRRNFVN
HLDSLRFSQNALITFHGGLFMPLIIGFFNQLVIQVPRVVIRPNTTAVYDLNOTSQEADDS
GDPLAIGOTINFLLCMILLVINTFFFVRSVRRNLHRRPHR

PAGE

CPMI0067 82920 84053

NO Tobust homolog present in Genebank/EMBL as of 11/7/98

KGSGYSYROPPMAVEGRVNSSOALNODCOEVLANKOSKGLLRCRILSIVVAVITPIAGVV,
LIALTLASILTSVPYLALGVFLLIVTLGCIIFALCSEKIKKVPPTPISHKEEIIAWFEBR

KMHDMEKEKEDPEHFGRTATDIPMRSALDOFINISCHHIHESPALTETYRSHODY LPKDW

CPVTLPDVTSEEEVLIRSVVGSYLLMEACVPKVSMLIDELHNUKLSPSERECLE PIKKTL

ORKASFIETOKOLLATFELAYTRVNIOGHLAPFRAGAWWILIHYVGLRRQHNONFFTHAGT

CYVARLAFNOTORLYHOLFNVEKLRSIYANMDKDPLCHPWAFIPIYDLLKTEDHGDGFLE

QGEDREYPSRAAQDOFWG

84909 84331 CMT_U068 84909 84331
CTJ#50 hypothetical protein
SF#IKKF1YSLIFSCSFSAPLKGICNEDVSSQSRIEEDPEVLITQLNELIETPIEEGKE
IR#ELQAISDCQKSSEEIEESCGTSDSEGLSEKTDKESSNEYVLDFFDSMVQRLEGISKM
CQSGQVAQIIDCFNREFDIRNRELELKNRELELREKDLEFKKSILDWYKEKVSRELAFQR EQDIKOTLMLLKK

CPI_0069 85191 87086

No robust homolog present in Genebank/EMBL as of 11/7/98

LNFLYVYLLIFNLGIMTTPPPSRSSSPPPYDWIELQDLGNTNNNSSRATPPPPEVGGELP
PYFSASNFVVIERGAPSLPSPQOLLSLPEVSRQPPPGYFDETASITSRTSEEMFGTLVST

LCCPANSERDWEDHEVNCIYIASTSDTOLEAVOGGMHITELRGEPPVRVLVETCHLYAFAR

ENTCHSRLEVSHTVRAMTYFWDRFFSRMWNVGRRFLVFYGGNGAYVQAALDSSMHTODI
YVLGLSPTVYIRGNYHVOHYRVRGFWPSCLDSLAACAENTSVLPYGGSSDGIFYPSLFSH
TFDNNAIRVGERCLLVCSEDGMGHPETQQOTSPLTSLEGGHEVALVLNPQONPEALSIASR
LMHEERCGRLESNYMPGRSSNFFMTSMYVLVRLNTLAQIYLMSPYYSFOSNDIVCLIFIS
SAAVETVGYIFLTVTDSTCGRRYLRVPRLVCTGLRNLALBTTLLELLILSYPRSVEGVPF
MVRFILGWTCTRVVFFANNLILHWFFRCLRRGIOLFVH/SIIGHTLGARITDLTLASMR
YAIVFPSIVSSCLLTALAHANTNILALDPYRLIESGDLHRPAFNDDEMQQADNPWDAYSI
GLVINTCIYMLILFANLIFMYSVRRYHRSRR 85191 87086

CPn_0070 87399 87208
No robust homolog present in Genebank/EMBL as of 11/7/98
YKVGLFHLKNONFFSNQSRTYEQRFPKVSPHFESII/PLQSVGFSSQGTLLISFRDTELKR

88066 87599 CPn_0071 88066 87599 / CP125 hypotherical protein LKGLRGTLEEFCPFGKDHLYLKLMENSSSRDAFDKKRMLT KERLEVEXQUOLYLVEVYQDGILFFFTYTKALMGSGIASLFTEVYSGETPSTILTCKPIF FORLTYYLFGRLMXHESLYMRMKGIAVOYLKPPQT

89151 88067 CITE 2007: METAL MADO 7
CITE 24 thypothetical protein
IGYESTETEVIKEKULLLIYCLLEYEFHYMMSTPLSSCGISPSDQYVPQELFCDRLSSSR
SNSFDENACIGST IVSPPISALVALTDLKLVPYNÇNSFSWTTRLKNAVEKIGLFLQRNWK
YILLYILAWALTIACHITVALTLTIWLAVGLSIGVIFGIFTATCLDKENKHRHVNSLWNL
INNGILDDEWSTRQILLATMIASISALIYAVPQAVGLVIGFSIGNQLSINTVYGARLGD

iq i i kiigm i nokolnal i e i nrnngtðpatanllas Lolmnnif do i i aradog i mtlsgtlog i kkepdr i EATYA LORKAHKKP LEN L LKLNLNOPMPYCFGMPEC(

CPn_0073 89353 infA-Initiation Factor IF-1 89574

SMAKKEDTLVLECK/EELLPCMHFRVTLENCMPVTAHLCGKMRMSNTRLLVGDRVTVEMS AYDLTKARVVYRAR

90955

TITE OF THE STATE VELDVELIGTVALEEGMRFAIREGGRTIGAGTISKINA

91087 91350 SREGF-preprotein translocase SRSWFMKQQHNRKALSRKIGTVKKQAKFAGSFLDEIKKIEWVSKHDLKKYIKVVLISIFG FGGAIYFVDLVLRKSITCLDGITTFLFG

n 0076 91334 91903 CPI_0076
71534
7150
PMsG-Transcriptional Antitermination
PPCSVNCMYKWYVVQVFTAQEKKVKKALEDFKESSGMTDFIQEIILPIEWVMEVKKGEH KVVEKYIWPGYLLVXMHLTDESWLYVKSTAGIVEFLGGGVPVALSEDEVRSILTDIEEKK SGVVQKHQFEVGSRVKINDGVFVNFIGMVSEVFHDKGRLSVMVSIFGRETRVDDLEEMOV **EEVAPGQESE**

CPn 0077 91956 92435 T111-L11 Ribosomal Protein
FFVSYPLFVEVSQCKVRFSMSVKKVIKIIKLQIPGGKANPAPPIGPALGAAGVNIMGFCK
EFNAATQDKPGDLLPVVITVYADKTFTFITKQPPVSSLIKKTLNLESGSKIPNRNKVGKL
TQAQVEAIAEQMHKMDIVLLESAKRMVEGTARSMSIDVE

CPn_0078 92453 93160 T11-L1 Ribosomal Protein
SCRIMTKHGKRIRGILKNYDFSKSYSLREAIDILKQCPPVRFDQTVDVSIKLGIDPKKSD
QQIRGAVFLPNCTGKTLRILVFASGNKVKEAVEAGADFMGSDDLVEKIKSGWLEFDVAVA
TPDMREVGKLGKVLGPRNLMPTPKTGTVTTDVAKAISELRKGKIEFKADRAGYCNYGVG KLSFESSQIKENIEALSSALIKAKPPAAKGQYLVSFTISSTMGPGISIDTRELMAS

93170 93688 Tilo-Lio Ridosomal Protein
RGNNKQERTILLQEVEDKISAAQGFILLRYLRFTAAYSREFRNSLSGVSAEFEVLKKRIF
FKAIEAAGLEVDCSDTDGHLGVVFSCGDPVSAAKQVLDFNKQHKDSLVFLAGRMDNASLS
GAEVEAVAKLPSLKELRQQVVGLFAAPMSQVVGINNSVLSGVISCVDQKAGKN

CPn_0080 93720 94121 r17-L7/L12 Ribosomal Protein vRVTKVTTESLETLVEKLSNLTVLELSQLKKLLEEKWDVTASAPVVAVAAGGGGEAPVAA EPTEFAVTLEDVPADKKIGVLKVVREVTGLALKEAKEMTEGLEKTVKEKTSKSDAEDTVK KLODAGAKASFKGL

CPn_0081 94219 98016

rpob-rna Polymerase Beca

freilshonsrrtrhlkcpervsvkkkedipdlpnlieiqiksykofloigklaeereni
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpiksyneatvleylsynlovpkyspeecirrgitysytlkvrfrldetg
gleevfreifpikofffiingaervvvsovhrspeinyfecekskonilfsfriipy
rdswleaifdindliyihidrkkrrkilaitfiralgyssdadiiefftigesslrse
kdfallvgriladniideasslvygkageklstahlkrhldagiasykiavdadehhpii
kmlakdpydsysealkdfyrrirreoepatlanarstimrlffpfrsynlorgvryklnrk
lofsiddealsovtlrkedvigalkylirlkwodekacvddidhlanravrsvgeliono
crsglarmkkivrermnlfdfssdtitpckvvsakglasvlkdffgrsolsofmodtnev
aelthkrrlsalgpgginreragfevrdvhashygricpietpegpniglitslssfaki
nefgfietpyrivrogivtdiefyrlydbeecviaoasasldeynmftepvcwwryage
affaddstythhdvspkolvsivtdiipflehddanralmgsnmorqavplikteapvvg
tglecraakdsgaivvabedgvvdfvdgykvvaakhpptikrtyhlkkflrsnsgttin
ogpleavgdvitkgdviadgratdrgelalgknvlvapmpwygymfedaiiiseklired
aytstyleffeltardfylkgeeitrdiphvsdevlandegoiirigaevkpolilvek
itpksetelapeerllraifoekaadvkdasltvppctegvvmdvkvfsrkdrlsksdde
lveeavhlkdlokgykngvatlkteyerkelgalllnekapaaiihrrtaeitvheglipd
oetierieoedlvblimencemyevlkgllsdyftalgreinykteveiriegdadldh
gvirovkvyvaskrkloyddmagrhgnkgvyskivpbadmpylsngetvomilmplgvp
srmligovlethlgyaaktagiyvktpvfegfpegriwdmiegclpedgksflydgktg
erfdnkvvigyiymlklshliadkiharsiopyslvtogoplogkaomgorgemewwal
eaygvahmloeiltvksddvsgrriyesivkgenllrsgtpesfnvlikemoglgldvr
pmvvda

CPI_0082 97992 102221

rdc-RNA Polymerase Beta'

CSSYGRRIKNDVLEKIMFGENSRDIGVLSKEGLFDKLEIGIASDITIRDKWSCGEIKKP

CSSYGRRIKNDVLEKIMFGENSRDIGVLSKEGLFDKLEIGIASDITIRDKWSCGEIKKP

ETINNRTFKPERGGLFCEKIFOPTKDWECCCCKYKKIKHKGIVCDRCGVEVTLSKVRRER

MAHIELAVPIVHIWFFKTPSRIGNVLGMTAGDLERVITYEEVVVIDPCKTDLTKKQLIN

DAOYREVVEKWCKDAPVAKMCGEAIYDLLKSEDLQSLLKDLKERLRKTKSQQARMKLAKR

LKIIEGFVSSSNHFEMMVLKNIPVJPDLRELVPLDCGFFATSDLNDLYRRVINRNNRLK

AILGREVPVSGRSVITVOPELKFROCGIFKENALELFEPFITIKRIKDQGSVYTIRSAKM

LORGAPEVMDVLEGIIKGHPVLLNFAPTLHBLGIDAFEPVLIEGKAIRIHPLVCAAFNAD
FORDÖMMAVIVPLOVEADLEANVLMMAPDNIFLEGOGFPVAIPSKOMTIGLYYLMADPTVF

PERHOSKTKIFKDEIEVLRALNNYFIDDVEGLBRDETTGGIHIHEKIKVRIDGOIITETT

FORVLFRRIVFRELJFOMYSMICKEIGLLIA/CYKKVGLEATVHRIJDLKDLGFIQATKA

ALSMGLKOVRIEDIKGHILKDAVJEVATIVKKVDGILTFRELDVENGEREIKTTGITUMFEVSGOLSD

ALYVELUKOTRSKHNPLFLMIDCHAPGNAKTADGOITTRELFOMYQDAFIDDADIFTRIK

GXGGEELLPLKORITYGRTVAEDVYGPIDKGHLAQGGOVINSVQARAIDDAGIFTIKIRIS

GXGGEELLPLKORITYGRTVAEDVYGPIDKGHLAQGGOVINSVQARAIDDAGIFTIKIRIS

TLTTGGJFRYGCAKCYGLNLANGLIGMERAVGIIAADGIGERTTGGTMRTFHILVILATS

STEFELTTREGGILVYMOLRVVILOVEZONNIGEVELHHIPIIGDKRGFFKYEDLVEGGTTEKVV

SLEVPPVELGVKILVADOTPVGXXXVIGEVELHHIPIIGDKRGFFKYEDLVEGGTTEKVV CPn 0082 97992 102221

NKNTGLVEL IVKQHRGELHPQ IA IYDDADLGELVGTY I SVEEGORVDPGMLLA RLPRGAIKTKDITGGLPRVAELVEARKPEDAADIAK
GMEEEHLIPLTKHLIVQRGDSVIKGQQLTDGLV/PHEI
GVRELQKYLVNEVQEVYR
LQGVDINDKHIEIIVRQMLQKVRITDRGDTTLLFGEDVNKKEFYEENRRTEEDGGKPAQA VPVLLGITKA;LOTESFISAASFODTTRVLTDAACCSKTDYLLGFKENVIMGHMIPGGTG FETHKRIKQYLEKEQEDLVFDFVSETECVC

102296 103312

TSILHOESDLIS

d Transport de la co quantiquaevamonoprogramation avois colorrocardos coloritarda alevagerak JELLNEAVVWG IRONGODLOTLSFILDK I OVNFALE I IKNI FOR I SLE I DARLSFNVEAM VORAVFLSOLFEAMCGDKKRLLVK I POTWEGIRAVEFLEAKGIACNVTLI FNLVOA I AAA KAKATLI SPFVOR I YDWWI AAYGDEGYSIDAD PGVASVSNI YAYYKKFGI PTO I MAASFR TKEQVLALAGCDLLTISPKLLDELKKSOHPVKKELDPAEAKKLDVQPIELTESFFRFLMN EDAMATEKLAEGIRIFAGDTQILETAITEFIKQIAABGA

103356 103751 predicted ferredoxin SEMKNKMDYKSQLVFSCPCCCKGNVCFSVFNLDVILTCNVCSSTYTFDSVIRNEIRQFVA LCKRIHDANSILGNATVSVSVEDNQMDIPFQLLFSRFPVVLNLSLDGKKIAIRFLFDALN

103766 104512 CPn 0085 CPH_0085

104512

101766

CT311 hypothetical protein
FSMKFFILFILIVAQFPAFSAQPRTQVSASHSKQAKARRTSRIRSSAATNASVSRYKTRA
AARKKIGKFEKKPSLSPVQWVRYSGKNYSIQTPSLWQCIDDKTQLPEKLDVLLIGKGKGN
LTPTINIAQEITSKSSKEYIEEILAVHKANEMTLESGIFTQIQSPSGEFTIKKTEKNSSW
GRVFCLQATTVIDHTAYIFTSTATLDDYAELSFTFLKVVSSFQIRGGKEATSGDAILEKA

CPn_0086 104898 atpE-ATP Synthase Subunit E 105527 ACDE-ATP SYNCHASE SUDDITE E
NIMANILANGKIKOJCOLARLDTIKPAEDEBAALLHNAKEQAKRIIQEAQEEARKILETA
EERAHOKIKOGEVALSQACKRALEALKQAVENKIFRESLVEWLEHVITDPEVSTKLIQAL
VQALEAQGVSGNLTAYIGKHVSFRAVNELLGKAVTTKLRKKSVVVGSFVGGVQLKVEEKN WVLDLSSSALLEIFTRYLQKDFREMIFQGS

105540 CPH_0087 105540 106376
CT309 hypothetical protein
SHEKIFSIFKVVMTQYYFLSSFLPTQLPESVPLFSISDLDDLLYLNLSENDLCNYGLLK
RFPDFENFAFFWAGKPIPFSFGEVTQENVERMLSSQOMSDDNDFEDFFKDFLMNHKSSQD
RENNFSDLFREFLSYHQTNSSKFLQDYFRFQQQLRVVLAGFRARVLNMDVSYVLRDEDSS
DENVLEVLMGKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DENVLEVLMGKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DENVLEVLMGKDSPNYELPEEFSDLGGVLDDYGLLPHTLNRALALYQFHKLEGFCSDSYF
DENVLEXTYMFAIRNSLASVEKGREIINHIEKAIKW

THE OOS 106352 108145

CT-288 hypothetical protein
SYRKCHOMYTYSEOTAOGHYIEAYGILLRVEFDGYVROGEVAYVNVDHTWLKAEVIEVAD
DELWYOVFEDTOGACRGLVFTSGHLLEAELGPGLLOGIFDGLORRLEVLAEDSSFLORG
KHTNAISDHNLWNYTPVASVGDTLRRGDLLGTVPEGRFTHKIMYPFSCFOEVTLTWYTSF
GTYNAHTVVAKARDAOGKECAFTMYORWPIKOAFIEGEKIPAHKIMDVGLRILDTWYEYL
KGGTFCTROPFGGGKTVLOHHLSKYAAVDIVILCACGERGEVVEVLOEPHLIDENGY
SKHRTCIICHTSSMPVAARESSIYLGVTIAEYYRQMGLDILLLADSTSRWAQALRETSG
RLEEIPGEEAFPAYLSSRIAAPYERGGATTKHOSSGSLTICGAVSPAGGMFEEPYTGST
LAVVGAFCGLSKARADARRYPSIDPLISWSKYLNOVOGILEEKVSGWGGAVKRAOFTKS
GSEIGKRWEVVGEEGVSMEDMEIYLKAELYDFCYLOONAFDPVDCYCPFERQIBLFSLIS
RIFDAKFVFDSPDDARSFFLELQSKIKTLNGLKFLSEEYHESKEVIVRLLEKYMVOMA cbb_0088

108111 109466 CT289 hypothetical protein LECWKKOWYKWRKDMOTIYTKITDIKGNLITVEAEGARLGELATITRSDCRSSYASVLRF LDE.WRKOWYKWRKUMOTITYKITDIKGNLITVEAEGARLGELATITRSDGRSSYASVLRF
DEKKYTLOVPGGTSGLSTGOHVTFLGRPMEVTFGSSLLGRRINGIGKPIDNEGECFGEPI
EİAIPTFNPVCRIVPRSMVRTNIPMIDVFNCLVKSQKIPIFSSSGEHKNALLMRIAAQTD
ADLĪVIGGMGLTFVDVSFFVEESKKLGFADKCVMFIHKAVDAPVEÇVLVPDMALACAEKF
AVĒEKKNUVLULIDMTAFADALKEISITMDQIPANRGYPGSLYSDLALRYEKAVEIADGG
SKĪĢITVTTMPSDDITHPVPDNTGYITEGGFYLRNNRIDPFGSLERLKQLVIGKVTREDH
GDLANALIRLYADSRKATERMAMGFKLSNWDKKLLAFSELFETRLMSLEVNIPLEEALDI
GUMVILASSPSSPFNGTARADILMEISDRAFICEV GWKILAQSFTSEEVGIKAQLINKYWPKACLSK

CPn_0090 109439 110080
atpD-ATP Synthase Subunit D
VLAKSMSVQVKLTKNSFRLEKOKLARLOTYLPTLKLKKALLQAEVQNAVKDAAECDKDYV
QAYERIYAFAELFSIPLCTDCVEKSFEIQSIDNDFENYAGVEVPIVREVTLFPASYSLLG
TPIWLOTMLSASKELVVKKVMAEVSKERLKILEEELFAVSIRVNLFEKKLIPETTKILKK IAVFLSDRSITDVGQVKMAKKKIELRKARGDECV

CPn_0091 110074 112053
atp1-ATP Synthase Subunit I
VRLNIHKYLFIGRNRADFFSASRELGVVEFISKKEFITTEOGHREVECLKVFDHLEAEYS
LEALEFVKDESVSVEDIVSEVLTLUKEIKGLETVKALRKEIVRVKPLGAFSSSEIAELS
RKTGISLREFFYRTHKDNEDLEEDSPNVFYLSTAYNFDYYLVLGVVDLEPDRYYTEIEAPRS
VNELGVDLANLGREIRNRSDRLCDLYAYFREVLRGLGNYNDEORLHGAKECCEDLFDGK
FAVAGWVIVDRIKELGSLCNRYQIYMENPVPDPDETIFTYLENKGVGMGEDLVQIYDTP
AYSDKDPSTWVFFAFVLFFSMIVNDAGKGLEFLMSSLLFSWKFRRKMKFSKHLSRMLMT
AILGLGCICWGTTTTSFFGMSFSKTSVFREYSMTHYLALKKAEYYLOMRPKAYKELTNEY
PSLKAIRDPKAFLLATEIGSAGIESYVVYDKFIONILMELALFIGVVHLSLGMLRYLRY
RYSGIGMILFWVSAVLYVPIVLGTYSLHYLFHVPYELGGOIGYVGMFGGIGLAVVLAMI
QREWRGVEEIISVIGVFSDVLSYZRIYALGLAGAMMGATFNOMGARLPMLLGSIVILLGH
SVHIILSIMGGVIHGLRLNFIENYHYSFDGGGRPLPPLRKIVCSEDAEASGIHLDNNSIV 110074 112053

CPH_0092 LLY121 LL2573

=SILDAK-ATECSYNEHOSE SIZDUNTE K

EYLKGAHEVSHILDMOVYGYALVLGLAMIGGA IGGSMAGVASHAVMSRIDEGHGKLIGMGA
MEGSGGI YGETIMLEMOYA IKNGTLSPVGGIA IGLCVGAALLVSSVMQGKGCVXGIOAYA
HOGGIYGKCYAA IGIVYDESLFAVVFALLLL

112440 113015

T1340 II 3915
CT103 hypothetical protein
CKACVVIAEEKIMADIROMGSIMORIGISHIFHCLLEFIRYYSKIVEGUTVILAA (SV
ICHA KORENIA KHEEKIVIEEKIAAN OLG LEGICHDEVYGOOVVITWILEGEMRKCLEVT
LYLWYYYGNGKYEKITYEVNOGAGYRVYCLKGLEYFELGG I GYRVALGGONGETVISRH

HUMMEV I SLOSP

CPn_0094 113104 11577

vals-valy1 erna synthetase
vyrvflsrdhkfglrimttedfpkaynfodtepel/vyfwekngmfkaeassdkppysvim
pppnytgvlhmchalvntlodvlvpykryggfevcwiptdhagiatqavverhloaseg
krrtdysredflkhtwawkeksek/vlsolrglgggdddrrrfteplanravkafkt
lefengy lyrgyylvnwdpvlotaladdeveyeekd wlly iryrmygggs ivvattrpe
tol. De a valtherpyscholal for sameraberth wiel depyskryftly
vilmthibei midilite valthe versekd processed before veryllownhlrdw
vivgyrsgaviepylskumpvov/fishabalbepylsolotkipfkdpvknyllomnhrdw
cisrglwyghriskympvov/fishabalbepylsolotkipfkdpvknyllomnhrdw
cisrglwyghriskkmpvov/fishabalbepylsolotkipfkdpvknyllomnhrdw
cisrglwyghriskkepytallvyrghollefwyrmyllcssmsgekpfseyflhglif
gksyrryndfgewsyiskkeklavomgealpdovakweklskskgnyldplemiatygt
davritlescanregoidldyrlfeeykhfankwngarfifghisdlogkbllagided
sleldfyildgffnglihgyreavatyafdkyatlayeffrndlespiggbolthagided
sleldfyildgffnglihgyreavatyafdkyatlayeffrndlespiggdaffghalr
mlrsracmeapypkafdykipqdresfflagrlvytirnirgenglddaffedgdaffchalr
mlrsracmeapypkafdykipqdresfflagrlvytirnirgenglddrespilkafvvcs
dtteigscipiloalgalesiglldkepekglysgvvdtirlgifypeehllkergrle
kervrleravenlerikgdesfcokanpnlvvakgealknnrielggildklasfa

CPn_0095 115956 118790
pknD-s/T Provein Kinase
ACIVCLDREDOSLERYDIVRIIGKGGMGEVYLAYDDVCSRKVALKKIREDLAENPLLKR
RFLREARIAADLIHPGVVPVYTIYSEKDPVYYTMPYIEGYTLKTLLKSVWQKESLSKELA
EKTSVGAFLS/FHKICCTIEYVHSRGILHRDLKPDNILLGLFSEAVILDWGAAVACGEEE
DLLDIDVSKEEVLSSRMTIPGRIVOTPDYMAPERLIGHPASKSTDIYALGVVLYQMLTLS
FPYRRKKGKKIVLDQQRIPSPQEVAPYREIPPFLSAVVMRHLAVDPQERYSSVTELKEDI
ESHLKGSFKWTLTTALPPKKSSSWKLNEPILLSKYFPHLEVSPASWYSLAISNIESFSEM
RLEYTICKKGLNEGGGILLPTSENALGGDFYQGYGFWLHIKERTLSVSLVMSSLEIQRG
QDLESOKETFLIALEQHNHSLSLFVDGTTWLHHMYLPSRSGRVAIIVRDMEDILEDIGI
FESSØSLRVSCLAVPDAFLAEKLYDRALVYRRIAESFFGRKEGYEARFRAGITVLEKAS
TDNMEGEFALAIEEFSKLHDGVAAPLEYLGKALVYQRIQEYWEEIKSLLLAKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOAITPGQEEKILLALKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOAITPGQEEKILLALKRYSQHPE
IFFLKDHVVYRLHESFYKRDRLALVFMILVLEIAPOAITPGQEEKILVWLKOKSRATLFC
LZDPTVLELRSSKMELFLSYWSGFIPHLNSLFHRAWDQSDVRALIEIFYVACDLHKWQFL
SCIDIFKESLEDOKATEEIVFSFEDLGAFLFAIQSIFNKEDAEKIFVSNDQLSFILLV
VIFDLFANRALLESGGEAIFQALDLIRSKVPENFYHDVLRNHEIRAHLWCRNEKALSTIF
ENYTEKQLUKEOHELFFLYGGYLALIQGAEAAKOHFDVCREDRIFPASLLARNYNRLGLP
KDALSYQERRLLLRQKFLYFHCLGNHDERDLCQTTYYHLLTEEFQL 115956

CPn_0096 124347 118837
CT296 hypothetical protein
ETFLSILREFFMKSLPVYVSGIKVRNLKNVSIHFNSEEIVLLTGVSGSGKSSIAFDTLYA
AGRKRYISTLPTFFATTITTLENPKVEELHGLSPTIAIKONHFSHYSHATVGSTTELFSH
LALLFTLEQAADPHKEVLDLYSKEKVLSTIMELSEGVQISILAPLLRKDIAAIHEYAQ
QGFTKVRCNGTIHPIYSFLTSGIPEDCSVDIVIDTLIKSENNIARLKVSLFTALEFGEGH
CSVLSDEELMTFSTKQQIDDVTYTPLTQQLFSPHALESRCSLCQGSGIFISIDNPLLIDE
NLSIKENCCSFAGNCSSVLYHTIYQALADALNFNLETPWKDLSPEVDNIFLRGKNNLVLP
VRLFQQTLGKKNLTYKVWRGVLNDIGDKVRYTTKPSRYLSKGMSAHSGSLCKGTGLGDYA
SVATWEGKTFTEFQOMSLNNMHVFFSKVKSPSLSIQEILQGLKQRLSFLIDLGLGYLTPN
RALATLSGGEQERTAIAKHLGGELFGITYILDEPSIGLHPQDTEKLIGVIKKLRDQGNTV
LIVEHEERMISLADRI IDIGRGAGIFGGEVLFNGKPEDFLNNSSSLTAKYLRGELTIPIE
ESREAPTSWLLITEAT INNLKNLSIRLPLARLIGGTGVSGSGKSSLINNTLVPAIESFLK
GENPRNLHFENGCIGRLIHITRDLFGRSQRSIPLTYIKAFDDIRELFASQPRSLRQGLTK
AHFSFNQPQGACIQCQGLGTMTISDDDTPIPCSECQGRYHSEVLEILYEGKNIADILGDH
TAYEAEKFFISHPKIHEKIHALCSLRLDYLPLGRPLSTLSGGEIGRLKLAHELLFFASPK 124347 CPn_0096 118837 AHFSFNOPGGACIQCGLITMTISDDDTPIPCSECQGKRYHSEVLEILYBCKNIADLDM
TAYEAEKFFISHPKIHEKIHALCSLRLDYLPLGRPLSTLSGGEIQRLKLAHELLFASPKO
TLYVLDEPTTGLHTHDIQALIEVLLSLTYLGHTVLVIEHMHVVKVCDYVLELGPBGGDL
GYLLASCTPKDLIQLNTPTAKALAPYIEGSLDIFVVKSEPPSSPRSCDILIKDAYQNNL
KHIDLALPRNSLIAIAGPGASGKHSLVFDILVASGNIAVAELFPPYIRQCLLKETPLPSV
GEVKGLSPNISVRKGSSSNRSYHTIASALGLSNGLEKLFAILGEPFSPLTEEKLSKTTPQ
TIIDSLLKSYKDDYVTITSPIPLGSDLEIFLQEKOREGFIKLYSEGNKVDLDERLPLNLI
EPAIVIQHTKVSPKNSSSLLSAISVAFSLSSEIWYIISOKKORKLSYSLGWKDKKGRLYP
EITHGLLSSDHPEGRGLTCGGRGEILKISLEEHKEKIAHYPLEFFSLFFPKSYMKPVQK
LLKDENASQPLKLLTTKEFLNFCRGSSEFPGMNALLMEQLDTESDSPLIKPLLALTSCPA
CKGSGLNDYANYVRINNTSLLDIYQEDATFLESFLNTIGTDDTRSIIQDLMNRLTFISKV
GLSYTITLORQOTLSDGENYNKHLAKKISINLTNIVYLFEEPLSGLHPODLPTIVOLLKE
LVANNNTVIATDRSCSLIPHADHAIFLGPGSGPQGGFLMDSDTEVCPSVDLHANVPOTEV
CPKAPLSISKANHTRGSDRTLKVNLSIHHIONLKVSABLHALVAIGGVSGSGKTSLLLEG
FKKQAELLIAKGTTTFSDLVVIDSHPIASSORSDISTYFDIAPSLAFAYASLTOAKAKHI
SSTHFSTNTKQQCSDCQGLGYQWIDRAFYALEKRPCPTCSGFRIQPLAQEVLYEGKHFG
ELLHTPIETVALRPFIKKIQKEKALLDIGLGYLPIOOKLSSLSVSEKTALKTAYFLYQ
TPETPTLFLIEDELFSSLDPIKKQHLEALLDIGLGYLPIOOKLSSLSVSEKTALKTAYFLYQ
GSGKQGKLLFSGSPKDIYASKDSLLKKYICNEELDS GSGKQGGKLLFSGSPKDIYASKDSLLKKYICNEELDS

CPD_0097 124549 126006
pyk-Pyruvata Kinase
pyk-Pyruvata Kinase
DSMITRTKI ICT IGPATNSPEMLAKLLDAGMNVARLNFSHGSHETHGQA IGFLKELREQK
RVPLAIMLDTKGPEIRLGNIPQPISVSGOGKLRLVSSDIDGSAEGGVSLVPKGIFPFVPE
GADVLIDOVY HAVVVSSEADSLELEFMNSGLLKSHKSLS IRGVDVALPFMTEKDI ADLK
FGVEONMDVVAASFVRYGEDIETMRKCLADLGNFKMPI I AKI ENRLGVENFSKI AKLADG
IMIARGDLG IELSVVEVPNLQKMMAKVSRETGHFCVTATQMLESMI RNVLPTRAEVSDIA
NAIYDGSSAVMLSGETASGAHPVAAVI IMRSVI LETEKNLGHDSFLKLDDSNSALQVSY
LSA IGLAGIQI AERADAKALIVYTESGGSPMFLSKYRPKFPI I AVTPSTSVYYRLALEMG
VYPMLTQESDRAVWRHQACIYGI EQGILGNYDRILVLSRGACMEETNNLTLTI I VNDILTG
SEFPET 124549 126006

CPn_0098 127494 126001
NO TODIST HOMOTOR PRESENT IN GENEROLIK/FEMBL ABOOT 11/77/98
LVOKKEHOTIKET LEAPLYYLVOG LIAL/RITTPEGETY/JUNGGEGELAFY LUDYRKTAL
TNLALAFPEKTEDERYK LAROLLOHLI ITTLEELIALEQUA'NI DKLITTLYTEGEN PKGES
SIEFY/GIEDLEETTRINJEKO'N LILFY/SIIQANWELDELY TYKNYBUTAFAKAT KNORESK
KIFALBEVEKGK IVPEKNOTQOG TEALBY SILVATIVOKOALLMENYTYPLEGUSAFTTE
PALLAYE/GEFV LAVING/ROAKGEFV/FE/DABL/YANK/GIMP/GIVAT LMIJOMM/FILEKG TA
SOPEZ/MM/HERWIKK KONY TKKYEZ/SIILI/PATOAV/SIEPELEKALAEGE/GOTTTHLAL
GNADIILEELOPOFPEY/LITARNOQUILALE/CYPATEDLTINI/JIIL/KHERKT/GC/AVY
CKEPLEKGLDIPOAPLKNELRIFY/GEBULK/DREPENEKVE/GEBULT/VE CPn 0098 126091 127494

127527 127805 177865 COMPOSITE TO THE CONTROL OF THE CONT

127882 129133 CPT_0100 12913 127892
CT011 hyporhetical protein
TT011 hyporhetical
TT0 TKTKETTKLYKKEW

CPn_0101 129986 129141
ybbp family hypothetical protein
PSTLCNFSQYTTOGPSKTMPFDITYYTTPLLEIILIWMLNYLLKFFWGTRAMDVVFGLL
AFLELFVLADKLHLPIIRRLMLHVVNIAAIVVFIIFQPEIRLALSRIRFHGKKFFIDTQE
GFVEQLAASIYQLSERGIGALVVLEWKDSFDEYLSFSSVKINATFSEELLETIFEPSSPL
HDGAVILRGGILAYARVVLPLAHDTTQLSRSMGTRHRAALGASQRSDALIITVSEENGSV
SLSRDGLLTRGVKIDRFKAVLRSILSPKEHKRKPLFSWIWKR

CPn_0102 130099 131466
cyda-Cytochrome Oxidase Subunit I
FY1QFWKFMDALILSRIQFGLF1TFHYLFVPLSMGLSMMLVIMEGLYLVTKKQIYKQMTW
FY1QFWKFMDALILSRIQFGLF1TFHYLFVPLSMGLSMMLVIMEGLYLVTKKQIYKQMTW
FWVG1FALTFVLGVVTG1MG1FSFGSNMANFSEYTGN1FGTLLGSEGVFAFFLESGFLGI
LLFGRHKVSKKMHFFSTCMVALGAHMSAFWIICANSWMQTFSGYEMVMHKCKLIPALTSF
WGVVFSFTT1DRF1HAVLGTWLSGVFLVISVSAYYLWKKRHHEFAKQGWKIGTICAVIVL
VLQLWSADVTARGVAKNQFAKLAAFEGIFKTEEYTP1WAFGYVDMEKERVIGLPIFGALS
FLVHRNIKKFPVTGLDQ1PRDEWFNVQAVFQLYHLMIMLWGVMVALTLISWSAYKGWRWAL
KPFFLVILTFSVLLPEICNECGWCAAEMGRQFWVVQGLLKTKDAVSPIVQANQIVQSLVI
FSLVFIALLTLFITVLCKKIKHGPEEENDLTEFEVK

131465 CPn_0103 131465 132511
cydB-Cytochrome Oxidase Subunit II
NRGIFHELSLTSLLPLLMYVILGVAVFAYSFODGFDLGLGAVYLKAKEDKERRILLNSIG
PVWDGNEVWLVIIVGGLFAGFPACYATLLSIFYMPIWTLVLLYIFRGCSLEFRSKSESVS
WKIFWDIIFICSGTAISFFLGTIVGNLILGLPLSPDTSYASLSWLLFFRPYAALCGAVVA
SAFAIHGSCFALMKTSDSLNARIAGOFPYILSSFLVFYVLFLGASLISIPKRFDAFPTYP
LILLIALTSCCCVAAKTSVSKKRYGYAFIYSTLNLLSLILSAATLTFPNILLSTVDPQY
SYTIYNSAVETKTLKSLLIIVLIGLPFIITYTCYIYRVFRGKTNFPSIY

CPT:0104 133884 132676
CT017 hypothetical protein
EKEMPLQISMLLIALGTAINSPATYAADSQSVSFPEQLPSSFTGEIKGNHVRMRLAPHT
EKEMPLQISMLLIALGTAINSPATYAADSQSVSFPEQLPSSFTGEIKGNHVRMRLAPHT
DCTIIRFESKGDLVAVIGESKDYYVISAPPGITGYVFRSFVLDNVVBGEQVNVRLEPSTS
APPLVBLRGTQIQPASQEPHGKWLEVVLPSQCVFYVAKNFVANKGPIELYTQREGKKI
AMDLINSALNFAHIELEKSLNEIDLEATYKKINLVQSEEFKDVPGIQGLIQKALEGIQKA
TERMSLESQNTSIASSQCSTPRVSSSSVTTSLLSRHIRKOTALKTAPLTQGRENLEYSLF
RIMASMQCMDHSEALTQEAFYRAEQKKKQVLAGVLEVYPHVVKNNPGDYLLKAQENTIA
FENGTSINLEQWLGKRVTVECLPRPNNHFAFPAYYVVGIKEAS 132676

cen_0105 134883 134029 CPn_0105 134883 134029
CT0166 hypothetical protein
VPFRKFSNONPHLLIYCKKEHLOWPOTAKIRFTPKIAMKVKINDQLICIPPFISARW
SOFFIESOEGENKDGSTLRLHLIDGKIISIPNLDGSIIDIAFQEHLLYLETSQSGKEDS
RDDBKLGVGVLMVVLQQITKGNDIQVLPKNLISPLFSGTNPIEAILQHTPEHKDHPDAPT
DVLEKMADVIRVLSGNNATLLPRPEPHCNCMGCJGTGRVNNEEDTLAVSDKDLTFRTWDIM
OSGDKLYVTNPLNPSDGFSVYLGPPIGCTCGEPNCEHIKAVLYT

136374

CPn_0106 phoH-ATPase 135073 phDH-ATPase
ER/ETOMKKTMVIDTSVFIYDPEALFSFENTRIIIPFPVIEELEAFGKFRDESAKNASRA
ER/ETOMKKTMVIDTSVFIYDPEALFSFENTRIIIPFPVIEELEAFGKFRDESAKNASRA
LSMTRLLLENAKTKVTDGVTLPSGSELRIEVAPLSNDDRRGKLLTLELLKIIAKREPMVF
VTKSLGRRVRAEALQIESRDYESKRFSFRSIVRGFRELOVSQEDIENFYKNGYLDLELDV
VKSEPNEYFPMSACENHFALGRYYVSEGKIIALKAMDKSVMGIKPLNTEQRCALDLLLRDD
VKL売中LIGOAGSGKTILALAAAMHKVFDKETYNKVLVSRFIVPMGRDIGFLFGLKEDKLM
HMMÖPIYDMMEVLFSINOMGNSSEALQALMDAKKLEMEALTYIRGSLPKAFIIIDEAQN
TLEHREIKTIISRAGKGTKIVLTGDPTQIDSLYFDENSNGLTYLVGKFHHLALYGHMFMTR
TERTSELAAAAATIL

CPn_0107 137321 136392
CT058 hypothetical protein_1
KKSPPPVTPKEIPTOPKPPIPORPEVSPTPTDHIVPGSIEASPILGKKPSPDSMVSPLEL
FHKMLLENWTPVEEPFPWPPAEKNOKIFAWALNOSKLIFVSTSGNIAQPRLVTDSMSMMI
VNAANRTMSRDGAGTNOVLSAAVSVDSWGLSGRPLNPERGTPLNEGECRACMMRNAGGS
NHTGKQCKHYLLQGLLGPKAVDHHNKSQAAFDRCKNAYLINCFSLAQTLGVTFLQTPJISS
GIYAPPENRKKPNSEENKVRMRWIHAVKCALVAAMQEFGNEPGNTDRRMLIVLTDLKTPA
ITDPKKKSHL 136392 137321

137887 137303 CTO18
KNLFHYKAILMSIFNEEVFIISHRHTPLGGTSTALRNTPLVNPLHRTNLORIASYIPIFS
TFIGIKTLKGISSLOYSMVLMTONFSSVCKTLPCPEIYEELPKVRKEAWLEIFGIKALYY
LVLGVIKIIKLIVRYLCPCCRPPEPREPONPLTPTPLDMGQQIDAIFSTPTPTPTGFKDPF LDDLLQEDKKKAPHL

CPn_0109 138646 141783

11es-Isoleucyl-trna Synthetase
ROMMTADEVCKNISFAKKEEQVLKFWKDNOIFEKSLONROGKTLYSFYDOPPFATGLPHYG
HLLACTIKDVWGRYATMDGYYVPRRROWDCHGVPVEYEVEKSLSLTARGATEDFGIASFN
EECRKIVFRYVHEMEYY INRIGRWVDFSSTWKTMDASFMESVWVFQSLYNVGGLYVCGTK
VVPFSTALGTPLSNFEASONYKEVDDPSLVVRMPLQNDSASLLWYTTPPMTLPSNMAIAV
GETLYYVRIQDKKGEOWILSQCOVGRWF:NPEEFVILESFSGKDLVGRTYEPPFTFFQS
KREEGAFRY IAASFYEESEGTGVVHMAPAFGEGDFLVCKENHVPLVCVEVDAMSFFTEEIP
QYCGY/IKHADKEI HRVLKKEGR IFYHGTVKHRYPFCWRTDTPL/YKAVNSWFVAVEKIK
DKMLPANS:HWYDEHIORARFCKMLEGARDWAIDRNRYWTPLYFIKMSADSEILLVVGSI
RELEELIFYTQTTDIHHHIE IODLAIKVDKRPFHRIE PRYFFCWFDGAMPYAGHHYPFENOK
ETEEAFPADFIABGLDYTRGWFYTLTVIGAILEDRPAFRNAIVNGIILAEDJRMSKRLN
NYCHPKYVLDTY/GADALRIYLLDHVVVKASADLRFGDGVGCGGVLAGFYAGHTAFTTPPDTY
WY LIMCGRRKFWEAEDDPPRHAAFSTLYEVLTWFCKVIAFFYPFLAEDIYQKLKLEKEPES
VHLCGFIQVEMDKILPDLEKRMHDIREIVGLAHERKKKKRVRQPLANFY/FGSKORLS CPn_0109 138646 141783

LLKTFEGLIAEELNVKNVII SETYTT/KPNFRMLGKKVGSKMKEVQKALGEPN DV:ICBHTDMYTARSSALESVILDCOLREPLIVE GIARELNKINTHRRNOGLET/SGIRS

CPN_0110 14:775 41827

LepB-Signal Peptidade I
LSYPSIPHKOHYSLNKSPHILRSTYK/LKSKKLAHSPADKKOLOELLEOLEEAIFEHDOE
LSYPSIPHKOHYSLNKSPHILRSTYK/LKSKKLAHSPADKKOLOELLEOLEEAIFEHDOE
LSYPSIPHKOHYSLNKSPHILRSTYK/LKSKKLAHSPADKKOLOELLEOLEEAIFEHDOE
LSYPSIPHKOHSULEVIN TERPHINITIER THE ALMAN ET BLOEL IDAGETKINGEL FOR
LSYPSIPH BLOEL TO THE ALMAN ET BLOEL THE ALMAN ET BLOEL IDAGETKINGEL FOR
GOKTIIDFKOPNOSYGRIFPOTÉMYGOFFDHKEMHODEPNKIKDEHLSPATYALFOM
GOKTIIDFKOPNOSYGRIFPOTÉMYGOFFDHKEMHODEPNKIKDSPALSPATYALE
GOKTIIDFKOPNOSYGRIFPOTÉMYGOFFUNTALITANISTIANAVAILLPKYPEGGYEY
YTLLPERKEHLHLIRNNLTTSEF/LVAQCAYKYHOFKINTSGIAKAYAILLPKYPEGGYEY
SKOEAYOIGFGEIRYKLKSSHFLTOLNDKOVIELFNCGINFSSIYNPVNPLQAPLENNY
SKOEAYOIGFGEIRYKLKSSHFLTOLNDKOVIELFNCGINFSSIYNPVNPLQAPLENNY
FFNOCHLYIMDSPUT IKNDPYLOKEVITSETEKGESSETOPYIAFVDKGLPPEDFKEFVE
PIHNFGIOVPKGHVLVLGLEWPMSADSREFGFVPMENLIGSPLCTFWPIGRMGRLTGVSA
PTTLSGYLVSGIALATGLSIGYVYYQKRRRLFPKKEEKNHKK 41827

CPn_0111 44761 143934
CT021 hypothetical protein
OLQNRYPIMPNDSSTYFERILOKYLMKKOGKTLFLFLFLSFLFSTAFSGLFASGTSSLRT
IQENIFLAKTGDYTVLSRGSQRTFVLVKSTTPKTVWIEIIHFPCIAHKERPSLEQASWKT
VHOLESPSOVFVVSJSSEGSOFFSLNTRTKSLEPVGKSTTVPAFLQIFDLPLSPAPANV
IKTKGKENKPWSPKVSFEGAPLTSISVNAWQGLWPKDRGPLSETGILMYFTQPDISVFPL
WVSIETPKGTSIVRAVDIGHGATSPYVYSLPDSKTQ 143934

CPn_0112 / 144743 145093
gatB-(Pet112) Glu tRNA Gln Amidotransferase (B Subunit)
DSDFGVVNMKKYTHPEYRQVLFVDSSTGYKFVCGSTYQSEKTEVFEGKEYPVCYVSVSSS
SHPFFTGSKKF/DAEGRVDKFLKRYSNVRQPAQQPQPEEDALPAAKGKKKVVTKKKK

CPn_0113 145329 146405

pfra-Peptide Chain Releasing Factor (RF-1)

gfmkkkvapytnmlaeveiktisnpeifsnskeysalskehsyllelknaydkilnlekvl

addkoalafekdpehvydlegdinfnskvelerlnkilesilvepdepddulvimelragt

ggeaalbygdcvrmyhlyasskgmkyevlsasesblkgykevymgisgrovkrildyea

gthrvorypetetogrvitsaitiavlpepseedtellinebolkidfrasggggkn

vtdsavgithlpygdvvdcodersohkikdkamgilkarirdaemokrhneasamgsaov

ggdrserirtynfsonrvtdhrigltlynldkvmegdldpittamvshayhqllehgn

CPH_07114 1463/7 14751
hemk/A/G specific methylase
wpPTsysNMEIKKAIQBCTAYLDYYGVPLSDCEALYILMDLLEVSSRAKLFDLVGISET
WLMSYRKRLALRGORCPTAYLMGAVSFLGLERVDSRVLIPRTETELLAEYIINYLLHSE
EIØTFYDICCGSGCLGLAIKKSCPHVEVVLSDVC-PQAVAVANENAKSMGLDVKILLGDLS
AFYTRPADAFVCNPPYLSFNEIHIDPEVRCYEPWKALVGGSTGLEFYQRIAQELPKIVT
SYGVGWLEIGSSQGESIKNIFSKHGIYGRLHQDLSGRDRIFFLEMDGRDPVSSGAYS

Pn_0115 147279 148622

fth-signal recognition particle GTPase
MINSLEGKLSSIFSFLVSSRRINEENISESIREVELALLDADVNYHVVKDFISKVKEKIL
GEETHKHVSPCOOFIRCHHEELVAFLSDGREEFTIOKTPSIILLCGLOGAGKTTTAAKLA
DVYLKNKKKKVLVVPCDLKRFAAVDOLKILVAOTKAEFYOSOENKPINVVVKALAYAKE
NGMOFVILDTAGRLNIDNELMEELTAIOKVSOANERLFYMNVAMGODVLATVOAFDOSLD
LKGVILSMTDGDARAGAVFSIKHVLGKPIKFECCERIQDLRSFDPOSMAERILCMGDTI
KFVKEMERY ISEEEDAELGKKLVTAAFTYEDTYKOMKAFRMGFDRAKLCMMGFFNAKP
SOKEIEDSEOOMKRTEAIILSMTPEERKELVELDMSRMKRIASGCGLTLGDVNOFRKOMS
OSKKFFKGMSKGKMEOVRKMSGGNOWR

Pn_0116 148592 148972
rs16-516 Ribosomal Protein
eknyrrksvalkirlroggrrnhvyyrlvladvesprdgkyiellgwydphssinyolks
erifywlergaolsskafalvkogapgvysallskoearklvvrkkrrayrorrstoree AAKDATK

CPn_0117 148983 150071
trmD-tRNA (guanine N-1)-Methyltransferase
TGMKIDILSLFRGYFGGPLQTSILGRAIKQRLLDVQLTNLRDFGLGKWKQVDDTPFSGGG
MLLMAEPVTSAIRSVRKENSKVIYLSPGGALLTAEKSRELAAASHLILLGCHYBGIDERA
IESEVDEEISIGDYVLTNGGIAALVLIDAVSRFIPGVLGNQESAERDSLENGLLEGPQYT
TNRDHFKCDKISVVLEVNKLKRAKNFYCKVFGLDAMSCENKFCLPHBGKTIFWLREVQAE
KKNIVTLSLSLDCACEEDFCYLLRRWELFGGKLLEKQADEHAVWALAQDLDGHAWIFSWH
BMK

CPn_0118 150075 150464
r119-L19 Ribosomal Protein
KKENFRWYIMVNLLKELEQEQCRNDLPEFHVGDTIRLATKISEGGKERVQVFQGTVMARR
CGGGGETVSLHRVAYGEGMEKSFLLNSPRIVSIEIVKRGKVARARLYYLRGKTGKAAKVK EFVGPRSSKK

150520. 151164 CPn 0119 CPT_0119 150520 151164
vnb8-Ribonuclease HII
LMNTSISEIQRFLSMIAFEKELVSEDFSVVAGIDEAGRGPLAGPVVASACILPKGKVFPG
VNDSKKLSFKQRAQVRDALMQDPEVCFGIGVISVERIDQVNILEATKEAMLQAISSLPIS
PDILLUDGLYLPHDIPCKNIIGODAKSASIAAASILAKEHRDDLMLQLHRLYPEYGFDRH
KGYGTSLHVEAIRRYGPSN'HRKSFSPIKOMCAIV

151125 151778 omk-CMP Kinase gmk-GMP KINASE BELFONKANVOYOMNKILVDSPFOPDHOKOOUKLETIGAPAGVOKTTLVRMLEQEFSSAF AUTIGVTTKKPREGEVTVKOYHFVSHEEFORLLDRQALLEWYFLFGECYGTSMLEIERIW SLOKHAVAVIDIQGALFIRSRMPGVSIFIAPPGOEELERRLASRGSEEGSORKERLEHSL TELAAANQFOYVT INDOLNÇAYRVLKS IF LAEEHRITT.

151769 CHOIST hypother test protein
EH MIKKORFTNEKLNKLFISSEFESLVIIVATKOAKTRIAKODVESSNVATETLVLLDREGT OPEFTEE LVVTA: PTVERKR: EHTNOPEKDE: AYTWODVK

CPN_0132 152062 153723
metG-Methionyl-tRNA Synthetase
CKVMPQKVLITSALPYANGPLHFGHIAGVYLPADVYARE 152062 CKVMPQKVLITSALPYANGPLHFGHIAGV/LPADVYARA GDDVLYICGSDEFGIAI TLNADREGL/YQEYVDMYHKLHKDTFEKLGFALDFFSRTTNFFHAELVQDFYSQLKASGL IENRISEQL/YSEQEQRFLADR/VECTCPRCGFDHARGDECQSCGADYEAIDLIGFKSKIS GVELVKKETEHSYFLLDRWKDALLSFIGCCYLPDHVRKFVVDYIEHVRSRAITRDLSWGI PVPDFPGKVFYWFDAPIGYISGTMEWAASQCNPDEWRFFWLEDGVEYVQFIGKDNLPFHSVVFPAMEL/JOKLDYKKVDALVVGEFYLLERRQFSKSEGNYVDMDKFLSSYSLDKLRYVL AATALEDAS GEFERBAFFER SLEEDAFFER CDDVLY (CGSDEFGIAI

CPn_0123 155975 153774

recD-Exodeoxyribonuclease V (Alpha Subunit)

NSMERICCYLEQILVENKDSGDITAYIKIPHNETTPILIKGKLEOPLELGSPIQIYGVWSH
SPSNTKYROJHSYDSPLLYEYRGVFHYLTSKLIKGIGPKIAEKIIEKFQEKTCYVLDITP

ERLSEVSGISETRCVSICKOLCEOKILRKTLLFLOEYNIPHYGVAIFKKYOEKSIEKIC

EDPFLLAREMEGIGFKTADFIAMKLGVPRNESSRICAGIOHSLEELGEECHTCYPIELLI

DVVAKLLNODVFDTPITLEEIDTQILNMQKRKLLHIQDISGTLHVWTRYLHLAEKTIVSD

LKRILFSSRRIRSIGGEKAIAWVEENLSIDLAEQOREAIKACFSEKLLIITGGPGTGKST

ITQAILKIFEOVTHKIILAAPTGKAAKRATEITGKHSVTIHALLQVDFKTKSFRKNHDNP

IDCDLIIVDESGMMDTHLLHHELKALPDYTTLVFIGDIHQLPSVGFGNILKDLITSNKMI

VIRLNKIFRQVHDSGIVTNAHRVNEGELPILVSETCRRDFLFFQKDDQEZALMHIHLMT

VIRLNKIFRQVHDSGIVTNAHRVNEGELPILVSETCRRDFLFFQKDDQEZALMHIHLMT

KFVPQKYHIYPODIOVLAPMKKGTLGITNLNKALKHALNPKKANLHGRFQSYAVGDKVMQ

IRNNYNKEVFNGDIGYVSTINFEDKAVVVRMEGKHVGYSFSELDDLVLAYATSVHKYQGS

ESPCIIIPHTSHFMMLYRNLLYTAITRGKKLVILVGTKKAIAIATRNNRVQHRCTGLAE

VLKELDTKKNYADL 155975 153774

CPn_0124 156575 158068

No robust homolog present in Genebank/EMBL as of 11/7/98
IRSKQRTVAITLLVLGILLIASGIIFLAVAIPGLSSAVALGLACGGMTALGTVLLITGLVL
LIRSEKLALEQVEIKQARTRVNNELDQLSQYVFYTENVLDNLKRWSYRDLGFVRQAQEEV
TNLEQDIEEIFITLRDIRNALDNEEFFMTHAKQCLAQVGESLFQDASIDEFINLAHLSEI
RQHLDINDPRWSMITKKVKGTVVRFIYVSTMYKQIKSNFEKSDFGQLRKMLLNNYMTIEE
VLYQSFQKGYNRAALLSERTRIHTSSLLHWEKDEDKHLNIKNECASRLENFKKFRTLFL
GLSEEDVIDFTGASGMCSKLPRKEVPLDGGKKKLAFKRTFADEQVGDMDRTTSLEHMTP
QEEDPLDRLMDQVEQEATSVLKDQDRYWKEIETSEAKFRSLPREDDFEKQSQIDSYIRDL
DDHLSWAANQLSAAEDALIEVTDVQEHGNREMLKNIQQGLELIEDAVKATLPRVDFIQEL
LEKEELPLVAARMSLENS

FTE-0125 158072 158605
No.Fobust homolog present in Genebank/EMBL as of 11/7/98
KISSCAEIMSEVKPLFLKNDSFDLATORFORLINHLOGOAEIYNEYYEEKNARVONEIKEQ
KOEFKLRG LEDPEARGIGUKKEELASITROFHDKAKAETSMLIECFCIGFYYSIHQEEQRQ
ROEFLQKMAERYRDCKQVLEAVQVEQKDMISSRVVVDDSYFEEEKEEQKVDNRKKEQD

CPR_0126 158806 161085
No.Exblust homolog present in Genebank/EMBL as of 11/7/98
No.Exblust homolog present in Genebank/EMBL as of 11/7/98
LLVFSYYCMGLFFFSCAISSCGLLVSLGYGLGLSVLGVLLLLLAGLLLFKIQSMLREVPK
APDLIDLEDASERLRVKASRSLASLPKEISQLESVIRSAAMDLAVIKTWHENKORLVEFV
SRKEERLAAAONYMISELCEISEILEEEBHHLILAQESLEWIGKSLFSTFLDMESFINLS
HLSTGRYLAVNDPRLLEITEESWEVVSHFINVTSAFKKAOILFKNEHSMKKKLESVO
ELLEFFIYKSLKRSYRELGCLSEKWRI IHDNELFBWVQDQOKYAHAKNEFGEIARCLEEF
EKTFFWLDEECAISYMDCMDFLNESIQNKKSRVDRDYISTKKIALKDRARTYAKVLDEN
PTTEGKIDLQDAQRAFERQSQEFYTLEHTETKVRLEALQCCFSDLREATNYRQVFTINSE
MANDLKSSFEKIDKERVRYQKEORLYWETIDRNEQELREEIGESLRLONRRKGYRAGYD
GRÉWELLRQWKKNLRDVEAHLEDATMDFEHEVSKSELCSVRARLEVLEELMDMSPKVAD
LEGLLSVEERCILDFIRENLERAYLQVMKCSEILSKAKFFFFEDEDELLVSEANLREVGAGL
KOVOSKCQERAQKFAIFEKHIOGOKSLIKEQVGSPLASKAFFFEDEDELLVSEANLREVGAGL
KOVOSKCQERAQKFAIFEKHIOGOKSLIKEQVGSPLASVGFLKSELLSIACNLYIKAVV
KESIPVDVPCMOLYYSYYEDNEAVVRNILLNMTERYONFKRSLNSIQFNGDVLLRDPYYQ
PEGHETRIKERELQETTLSCKKLKVAQDRLSELESRLSRR PEGHETRLKERELQETTLSCKKLKVAQDRLSELESRLSRR

162152 161130 VtlF*Cationic Amino Acid Transporter
ESFM*FPSANOESRTRNVPLGIFHGLVACLYWGIVFVIPNFLGSFGDLDIVLTF ESFMFFSANQESRTRNVPLGIFHGLVACLYWGIVVVIENFLSSEGDLUVUTHYTTESIF SLLÄGAIKHPSVIKKTPLVIMRKSLLUMTLLIMPVYFFGITLGIRVVGSAITVIISLAPT AVLYHSNTKQKELPYSLLFAISSVIITGVILTHLSALNLPTAASPLYSILGVIAVILSTS LWVIYVIRNQSLLEKHPNLTPDTWSYLIGISALIICLPMIIILDLCGITHYTHNLISHTP GSERLLFLLLCSAMGIFSSAKALIAMNRASLNLSPALLGAILIFEPIFGLVLTYLYSQSL PSLQEGIGIFLMLGGSLLCLVLFGRKVQKSLENSQVSSSNE

162262 163053 CPH_0128 16222 163033 bpl1-Biotin Protein Ligase EDRGRMLRNQVLVYCSEGVSPYYLRHTIRFLKYYSTQEGAFDILAVDGNFLIKNPFWEET TRLLVFPGGADRFYHRVLHGLGTARIFQXVSEGGNFLGTAGAFFGSKMIYFYEPEGAPL GGARDLGFFFGTAKGFAYKGNFSYYSPGVFRVSPQLFSDFGLGYAMFNGGCFFEGSEGYP GVNIESRYDDLPGKPASIVSRUSKGLAVLSGPHIEYLPHYYRMVKENVQKTREFLQRER TTLDRYCONLVORLROPAFSKADC

CPn 0129 163747 163064 CPD_0129
similarity to CT036
similarity to CT036
DEOYILSHIMMDPRIPYTSEPLOKTYQKLOEKHVNNLJIASOVSLTDLONKTQYENNLIE
TTTNEITYYFPVVHNPDILRSEWDPISNOLYLIFKKFIHYHNLFSTALERNOILLIDSL
NTGSSNPIAROMELLAFLCVFEQLDYNEDEYTIEPBV/FNRFVYKNSOTAPQIOSFGLLH
GYEEMSYASNNIRNVLTHSIVLCSPILYQLITEFDTTKIHADDFDCLI

CPN_0130 164251 163751
No robust homolog present in Genebank/EMBL as of 11/7/98
SEMANCES ETHENKERAGLEPESKFAATEKLIJATLESFLOTAACILITALSGLEPNTLLI
TALGLIGI INJERGISLI GTOGGSKOVOK DEŠKRE, IFPKERPSLEPSKLAPILAPILNELKNK IQSS
ETILLIDPESINLKNELFFFDFEEWKKIFLKPPDFLIKSALANWKILE

No robust homotost present by Genelonk/EMBL as of U177798 essekkiskijostalelijefftsayvees icelelemenamissevynipswilktisvaq evekkiskijovilstsymletiolsyvaeteegyllyfystialimla silvletilissy RESMYDRIAM TRESTYALIGHERGEPEZDYKRYZJÓ LULRISTY LKYRALMERGUT PEDPSOAA VILLIGHAPPERGEYTVERALLIGHERGEPEZKY LOFYLLAF LERVIGLYPEJARFILDOLMEQ GYMFFERINFERELFET KAKARERG ZODAKHE IMPGIEBETYGYPEJGEMERGYZJAMFISYYRY LROPDISTERLIGHELGEFRAGYZHOLAGEENPKDI ADGDFLEACKNVEWSEF LSACEK ALLKNPQGISIKDLKQFLVF

CPn_0132 165564 166561
No robust homolog present in Genebank/EMBL as of 11/7/98
SMIEFAFVPHTSVTADRIEDRMACRMINLSTLA FTSLCVLISSVCIMIGILCISGTVGTY
AFVVGI IFSVLALVACVFFLYFFYFSSEFKG/SSGEFRFLP IPAVVSALRSVEYISQDA
INDVIKOTMOLSTLSSLLDPEAFFLEFPYFNSLIVNHSMKEADRLSREAFLILLGEITWK
DCETKILEMIADPNTTPDDFWKLLKDHFDLKOFKKRIATWIRKAYPEIRLPKKHCLDKSI
VKKCCHPLLLUSBURV/CYGSILBKC/CYFSGEFPAMVLXILGGEVPMVLGLPKVPVDLTWEMF
MEITMI VILLOPPHYBURF LEGIVALL

CPT_0133 167349 66564
CHLPS hypothetical protein
NSSAYMFKLLKNIFLIGGCIVGYFMARKESIVEOWLSNRLHTQVTVGRVSIRTSGIKIRH
ICIHNPLASERFPYAAEIEYADVRSSISMLLTKOLEISELIIHGANFTIFPYDSHGTKT
WMSLVWKNFHPOKETPSNLWHDRAFVLIRRCLFLNTRLYGLRANHKDIPHLSVPSLEFHS
HTSSAKELPKLSEALPSLLYLALEESLYHLNLPGDIIKPLSQAHKHFYSSYPOFODRLN
DINTPGTPTEEIIGFIRGLFFH 167349 66564

167467 CPn_0134 groEL-HSP-60 169131

GTOEL-HSP-60

FADYRKIRRTTMAAKNIK/NEEARKKIHKGVKTLAEAVKVTLGPKGRHVVIDKSPGSPQV
TKDGVTVAKEIELEDKHE/MGAQMVKEVASKTADKAGDGTTTATVLAEAIYSEGLRNVTA
GANPHDLKRGIDKAVKVV/DELKKISKEVQHHKEIAQVATISANDSEIGNIJAEANEKV
GKMGSITVEEAKGFET/LDVVEGMFINRGYLSSYFSTNPETQECVLEDALILIYDKKISG
IKDFLPVLQVAESGRPLLIIAEEIEGEALATLVVNRLRAGFRVCAVKAFGFGDRRKAML
EDIAILTGGQLVSEFLGKKLENTTLAMLGKAKKVIVTKEDTTIVEGLGNKPDIQARCDNI
KKQIEDSTSDYDKJEKLQERLAKLSGOVAVIRVGAATEIEMKEKKDRVDDAQHATIAAVEE
GILPGGTALVROTPTLEAFLPMLANEDEAIGTRIILKALTAPLKQIASNAGKEGAIICQ
QVLARSANEGYDALRDAYTDMIDAGILDPTKVTRSALESAASIAGLLLTTEALIADIPEE
KSSSAPAMPSACMDY

169448 169143 CPn_0135

groes-10/Da Chaperonin MSDQATTVRIKPLGDRILVKREEEEATARGGIILPDTAKKKQDRAEVLVLGTGKRTDDGT LLPFEVOVGDIILMDKYAGQEITIDDEEYVILQSSEIMAVLK

CPn_0/136 171419 169569 CPD_0.36 171419 169569
pepF-0.1 igopept idase
KGVPSLMITELKTEALPTRTQVDPKHCWDTTLMYANREEWKKDFDLCSSGKDRSPIWPEF
SPSHYQIDNPESLLELLSKKFSVERKLDQLYIYAHLIHDQDITNPEGESDYQSIVYLYTL
FSOEISWIQPALIALSEEKVAALLSSSVLAPYRFYLEKIFRLSPHTGTANDERILASFI
ALNYSNKAFSSLSDAEIFFGIAKDSNGEEHPLSHALASLYMQSPDQELRRTAYLAQFQRY
/DYRNTFANLLNGKVQAHLFEAKARNYPSCLEASLFQRNIPTTVYINLINETKKHTSLIN
KYFNLKKEALNLKEFHFYDVYAPISQTTSKNYSTEGVDLVCKSLLPLGTHYVEILRNGL
LSNRWVDRYENKHRRSGAYSSGCYDSAPYILLNYTNTLYDVSVIAHEAGHSMISYFSREA
QPYHDAQYPLFLAEIASTFNEMLLMEALSKSDQSKEDKIVIITKTLDTIFATLFRQFFFA
AFEVEIHSAAEQGTPLTEEFLSATYGNLQKEFYGGVYTSDSUSSALEWARIPHFYYNFYV
QYATGIIAALSFAEKILTQEPGALELYLKFLKSGRSDFPLNILKKSGLDMTTSAPLDKAF
AFITKKIDLLSSLLSED

172263 171502 CPn_0137 CPH_U13/
ybg1-ACR family
vCSMNVADLLSHLETLLSSKIFQDYGPNGLQVGDPQTPVKKIAVAVTADLETIKQAVAAE
ANVLIVHHGIFWKGMPYPITGMIHKRIQLLIEHNIQLIAYHLPLDAHPTLGNNWRVALDL
NWHDLKPFGSSLPYLGVGGSFSPIDIDSFIDLLSQYYQAPLKGSALGGPSRVSSAALISG
GAYRELSSAATSQVDCFITGNFDEPAWSTALESNINFLAFGHTATEKVGPKSLAEHLKSE **FPISTTFIDTANPE**

174094 172700 CPn_0138 174094 172700

"hemL-Glutamate-1-semialdehyde-2,1-aminomutase"
TNSRLFLAIKDOLLONMMKLTKRNSMLNCSNOKHTVTFEEACQVFPGGVNSPVRACRSVG
VTPPIVSSAQGDIFLDTHGREFIDFCGGWGALIHGHSHPKIVKAIQKTALKGTSYGLTSE
EEILFATMLLSSLKLKEHKIRFVSSGTEATMTAVRLARGITNRSIIKFIGGYHGHADTL
LGGISTTEETIDNLTSLIHTPSPHSLLISLPYNNSQILHHVMEALGPQVAGIIFEPICAN
MGIVLPKAEFLDDIIELCKRRGSLSIMDEVVTGFRVAFGAQDIFNLSPDITIYGKILGG
GLPAAALVGHRSILDHLMPECTIFOAGTMSGNFLAMATGHAATAOLAGSEGFYPOHLSQLE
LFYSPIEEEIRSQGFPVSLVHQGTMFSLFFTESAPTNFDEAKNSDVEKFQTFYSEVFDNG
VYLSPSPLEANFISSAHTEENLTYAQNIIIDSLIKIFDSSAQRFF CPn 0138

174686 174093

yqge SPTKNKLRDIMKIPYARLEKGSLLVASPDINQGVFARSVILLCEHSLNGSFGLILNKTLG FEISDDIFTFEKVSNHNIRFCMGGPLQANOMMLLHSCSEIPEQTLEICPSVYLGGDLPFL OEIASSESGPEINLCFGYSGWQAGQLEKEFLSNDWFLAPGNKDYVFYSEPEDLWALVLKD LCCKYASLSTVPDNLLLN

CPn_0140 yqdE 175140 174673

yddi Prsnogkifomslekelleetpluvlnfyklusfonyagmilgteekkfaiyghusmoga Fogadteghsporpfahdllnfufsgfdigvlrvuindykdnufytrlfleokdreflyu VDVDARPSDSIPLALTHKIPILCVKSUFDAVUPYEE

175817 175110

CPD_0141 17817 175110 rpia-Ribose-5-P Isomerase H HSSSAVENDLHLHEKKCLAHEAATOVTSGMILGLGSGSTAKEFIFALAHRIQTESLAVHA HASSONSYALAKOLAIPLLNPEKFSSLDLTVDGADEVDPQLRMIKGGGGAIFREKILLRA AKRSIIL/DESKLVEVLCKFRVPLEISFFGRSAIIEEIRHLGYEGEWRLODTGDLFITDS SNYTYDTESPNSYPNPEKDLLKLTQTHGVTEVGFVTEKVEVWSSNSQGLTSKKYSV

176121 175816 No robust homotor present in Genebank/EMBL us of 11/7/98 SHSYSYSYSTEEKFHEKILOLLSTKNSTLNFSSHFEISRVSHDNAIQKIRSYPLKPIAEN RINTLOFFOLKIDYPKDSSKREPELYCIGRIPLHFLWEYFYT

177347 176214 **yxjG_Es_1 Hypothetical Protein
PRTHMTTSLKRPLKSHFDVVGSFLRPEHLKKTRESLKEGSISLDOLMGIEDIAIQDLIKK GARKATA DENADYCHANTATIKANINA BAKCONAG I DERGINDET I ODATE I NUEN I VO OKWANI POL GLOMERLANGATIKANI HANAGATANI LEMIN LEMAKKATANO ETI EDI NY OKWANI POL GLOMERLANGATIKANI HANAGATANI PARAMI DILANTAK I MARAMI RPDDLY/NLIIVORGNYHSKFFASGSYDFTAKPLFEQTN SGEKTVOLGLYTSKTPTLENKDEVTARTHQAADYLPLE EQWAKVALVKE I SEEVWK

efdhersgdfspltfi Ocgfasceignkltee

[77942 180560

CPD_0144 (77742 1800b0
clpB-Clp Proteise ATPASe
KVLCVIPMEKFSDAVSEALEKAFELAKGSKHTY/TENHLLLALLENTESLFYLVIKDIHG
KVLCVIPMEKFSDAVSEALEKAFELAKGSKHTY/TENHLLLALLENTESLFYLVIKDIHG
MPGLIATRAVKDALGREPTVYEGEVDREPSPGLOTLLEDAKQEAKTLGDEYISGDHLLLAF
ALGREGEFICHE FOTTYGEVELERIGGET FOTTGARGER, GLEYYTKILITALARE
GRELE VIT BEREETERPT FOVLUMETERIGH MIG GER CT SYTATARG ALBERTYTKILITALARE
GRELE VIT BEREETERPT FOVLUMETERFER FOT SYTATARG ALBERTY KAN KOMEGODENE FETTORIUM VARGARDAMO

KQLYVLDMGAL IAGAKYROEFEERLKSVLKDVESCOGERI II FIDEVITHVOAGATDCAMD
AANLLKPALARGTLHCIGATTLNEYGKYI EKDAALERRFOP IFVTEPSLEDAVFILKGIAMD
AANLLKPALARGTLHCIGATTLNEYGKYI EKDAALERRFOP IFVTEPSLEDAVFILKGIAEKYEI FROVENITEGALNAAVLLSYRI PDRFLPDKA IDLIDEAASLIRMQIGSLPLPIDE
KERELAAL IVKQEAIKREOSPSYQEEADAMOKSIDALREELASLIKJOEEKKLISGLKE
KKNSLESMKFSEEBAERVADYNRVABLRYSLI POLEBEIKQDEASLNORDNRILLGEVDE
RLIAQVVANNTG IPVOKMLEGEAEKLLILEESLEERVVGOPFAVSAVSDSIRAARVGIND
PORPLGYFLFIGFTGVGKTELAKALADLLFINKEEAMVRFDMSEYMEKMSISKLIGSSFGY
VGYEEGGSLSFALRRRPYSVVLFDEIEKADKEVAINLLLQVFDDGILTIGKKRKVNCKNAL
FIMTSNIGSPELADYCSKKGSELTKEAILSVVSPVLKRYLSPEFMNRIDEILPFVPLITKE
DIVKIVGIONRRIAQRIKARRINLSWDDSVILFLSEQGYDSAFGARPLKRLIQQKVVILL
SKALLKGDIKPDTSIELITMAKEVLVFKKVETPS SKALLKGDIKPDTSIELTMAKEVLVFKKVETPS

180717 182369

CPn_0145 180717 182369
CT114 hypothetical protein
NCASFIMLNKSSHRNLRSPMFKSFTVRYMFVGGLVSFLLPIPDLECANNVTKTYDKKAS
VISRDLKLQEDCQKFWNLDPYKLESLCAYQVLYHDDYSSKRIRELFPQIQKDEVPIFATM
ILTIGKVDRGFSPEEISLIQKLSYPGLSLASLRGSTEIDPNTDLARALVVSEFSGDLGKN
RADTYSNCLDILALRIHAERQRYLDQSPC-VPGTSEFHKATIEAINTILFYEEAVRYPSKK
EMFSDEFSFLSSVTDRKFGVCLGVSSLYFSLSQRLDLPLEAVTPPGHTYLRYQGGEVNIE
TTAGGRHLPTASVCDCLDLEDLQVRTPEEMIGLTFMNQGSFALQKKYKKEEAEXYKKAQE
YLGDEELQEGLIGFVQILGGKKKEGKSLIGKSPRASQKGSVAYDYLKGRINIPTLALLFSY
PGSNYEEIASYEEELKKAMKSSMPCCEGGRRLASVAFHLGKTAEAVALLEKCVEDIPNDL
SLHLRLCKILCDRHETTKALKYFIIAERLMEDQGFLKKDNRSFALFYEVKKIISKVAPQK
ANTLLLMESER ANTLLLMESER

CPn_0146 182595 183095
No robust homolog present in Genebank/EMBL as of 11/7/98
IIVGISMSSSEVVFOTVHGLGFGGLSSKSVVPFKKSLSDAPRVVCSILVLTLGLGALVCG
IAITCWCVPGVILMGGICAIVLGAISLALSLFWLWGLFSNCCGSKRVLPGEGLLRDKLLD
GGRSRAAPSGMGLPGDGSPRASTPSCLEELQAEIQAVTQAIDQMSDD

CPT-0147 183213 183671 No Frobust homolog present in Genebank/EMBL as of 11/7/98 HGGEMAVQSIKEAVTSAATSVGCVNCSSREAIPAFNTEERATSIARSVIAAIIAVVAISLL GUSTJVVLAGCCPLGMAAGAITMLLGVALLAWAILITLRLLNIPKAEIPSPGNNGEPNERN 183671 SATEPLECGVAGEAGRGGGSPLTQLDLNSGAGS

183822

CHIO148 183822 185702

pkni-s/t Protein Kinase
gkwirvsymesekdigakfigdyrilyrkgqslwsedllaehrfikkrylirlllpdlgs
sqørmeafhdvvyklaklnhregilsienysesegreflytqeqdipilslyqylksiprk
lifeleiudivsqlaslldyyhseglaqeemilbyvihtlirlgfasliker
lidefisdeenreskikervllhtsegkqgredtyafgaityyllfeglpgiffasliker
sgonlepaukkveetryshqokesaehlefylveacsideandtaiesessegveedys
Lalqsllvrepyvsryveaekeepkropiltemvilegefsrgsvegqrdelpakytl
sysllydhpytnegtfyleccsegddkyynelilridstirgsgriviegyakhfyv
gyfydygasgyaewigkrlpteaeweiaasggvaalrypcgeeieksranfftadttyms
tyrmpygasgyaewigkrlpteaeweiaasggvaalrypcgeeieksranfftadttyms
tyrmpygasgyaewigkrlpteaeweiaasggvaalrypcgeeieksranfftadttyms
tyrmpygarnypgropyryferakynin LRCAHRHRNNPGAVNSTYGFRCAKNIN

187700 185706 CPn_0149

CPD_0149 185706 187700
draft-DNA Ligase
ERSMKENSQAHYLALCRELEDHDYSYYVLHRPRISDYEYDMKLRKLLEIERSHPEWKVL
WSTSTRIGDRPSGTFSVVSHKEPMLSIANSYSKEELSEFFSRVEKSLGTSFRYTVELKID
GIAVAIRYEDRVLVQALSRGNGKOGEDITSNIRTIRSLPELLPEDAPFIEVRGGVFFSY
STFGIINEKOQOLEKTIFANPRNAAGGTLKLLSPQEVAKRKLEISIYNLIAPGDNDSHYE
NLQRCLEWGFPVSGKPRLCSTPEEVISVLKTIETERASLPMEIDDAVIKVDSLASQRVLG
ATGKHYRMALAYKYAPEAFATLLEDILVQVGRTGVLTPVAKLTPVLLSGSLVSRASLYNE
DEIHRKDIRIODTVCVAKGGEVIPKVVRVCREKREGSSEVMMPEFCPVCHSHVVREDR
VSVRCVNPECVAGAIEKIRFFVGRGALNIDHLGVKVITKLFELGLVHTCADLFQLTTEDL
MQIPGIBERSARNILESIEDAKHVDLDRFLVALGIPLIGIGVATVLAGHFETLDRVISAT
FEELLSLEGIGEKVAHAIAEYFSDSTHLMEIKKMQDLGVCISPYHKSGSTCFGRAFVITG
TLEGMSRLDAETAIRNCGGKVGSSVSKQTDVVVMGNNPGSKLEKARKLGVSILDQEAFTN
LIHLE

CPn_0150 187759 192444

CT147 hypothetical protein
IIYYKFFYSYNCPYFISFFVLLGVMMASSSNNSTKGDGIPSWVNPNVGWNRASQVGDEA
MSLTPEAQTSRSWFSDRKHFLEVLDVSLEEMENNDLKKYSRYKTIILIATLVTVAITCIV
PISWYFGIPMWVECLLFOAGLSSAFLSHRLQSKCKEHLHRYRAYQIYRQQLLSQYPDLR
KSTLYKYSITHVKPKKGFVGKLVENLRPDLHKNKDDGGAAADSRLDFAGYGVKHYWTDAL
LGVSGVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVYSGKDIGGEIDPGGIL
UGSVSCVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVYSGKDIGGEIDPGGIL
UGSVSCVNSVEWQRLASLIMSVKNDILNDVGSREPIDPAQRSALVYSGKDIGGEIDPGGIL
UGSVSCVNSVEWQRLASLIMSVKNDILNDVGSKEPIDAQRSALLYSGKDIGGEIDPGGIL
GYARKIFQDLSALTTAHGTGQALEDLDSLLCYYDQLIESKGVGEKIIASIHOKHUDLAMQD
PCDQEHLKKWSNLYHVFSITIKEFTEGKLEQNEVVSRIGNLERGKLESKKSILIANCETNA
EYATKSEKKLADYLLQIGDREPFLTGMHKAIATGKAIQGKVEGVISQHPEKOIMMLRGSI
EPLEGGMLRREDWGAILQKNEDEVLALKSTMEAQLOGFFLOVTTEGKLKSKVL
VYOFTKINSYSULLARLEVLHAESSTHODLVLHVDRMSEDLKKTIEEIDONLFQVTFEELSLL
AFEYQLIMNELPLIVOEGNRLQEAISSEOVSGILMLIIGLIANDEKINKNYESSRKNLVA
IAKQAR:DARNID:GGLAPLIQRNRASLDNILQMYZLFNGSIRNIHALDTETLVATSSMM LAVQAR: DAKN LO: COLAPLIORNRAS LONI LQNMY LENGS TRITHALDYETLVATSSIM
FSAMITTEDWITYTHLLDVLEIGSK PAPAPMEN PLLP GALPEEV QUAVAE DVSCTHRLHIQ
VLPTRI: ADI.KNM D: QUQKS INKWGMAKA LVLG IVAYLE CVLISA I FI CON I LISLLI LISCVO
LLL COVETA, FEDRI GSKSKEPER QVLETAGSL I PATP I LEGERNIND LA BLAKLQONLINLE
GPGTTMARN LV: GDLEG I TYKEKGLKDLTKEFRKDCY ILLNKR I KRREK GLIQQEAPVVP PT
LPJOL IR GREVPARLIRELEHLOKOKEET I TRODALY JEPMOLCLEKSKY OMEKAHAAAMT
KKYVKILON I DRI JKKNNETY VRI ONFRITL I GEKLOF LITVOET DVVKEAKELHELAATIYG
NTOTROKOKRAKKOPKENVLH I MOKOQLELLEAYLI TYTAGO ILCRI JONGAS FRER I LLIN
LOARRIGEARETI JAGREPMLKT LGLOYLITEFVRESDE ESTO: GYNNYLLKVRROLFDI EORL
ONGETVEHENVAAVQAALAAYVRKHEGLI VSTYGLSAMET I DVVKKAKEFELLNINGSNNPNDPEC QKYMOILLDAPVSLLYGAF; FILMETSUM IANGTKAAEEEAKRYVEEKGRGFETY EQCIPLAVILKISIFSDLMLREKVSVEKAALEEEIQ GIQEQYAEMQGIEDLELKQKFEDLOKKLEALEELLQIGRRIDSSVDKQKELLGLLGREE

625 194179 1.9. CPn_0151

CPT_0151
mhpA-Monooxygenase
CYENLERYPRASMADILVIGANPTGLILANMLIQHGI3VKVIDHRASPEDPSFLDCRKLP
CYENLERYPRASMADILVIGANPTGLILANMLIQHGI3VKVIDHRASPEDPSFLDCRKLP
CYELEVIGHEELEMADEX TOF FOANIE TO ARREWERP TGLEFFCOATCOPYPESIGTTYQ
CHECHLICHER PROFIT OF HOTHOUTH TO ARREWERP ARREWERP TO ARREWERP TRIPLET TO ARREWERP TO ARREWERP TRIPLET TO ARREWERP TO ARREWER HLEGHLIBER FROM THE THE THE TABLE OF THE THEORY OF THE TRIPLATIVE
CPn_0152 1952/4 194318
CT149 hypothetical protein
LIMMRVAFLVSCLFSVAIGASAAPVRVBGFPOIPEDLVQIKTEVCPKQEVCLAVTIKCD
DHNLIGVLHLPNTPTPEGGFPTVVLFHGFROTKFGGLTGAYRKLGRKFAAVGIATLRVDM
AGCGDSEGVAEEVPIETYLBDAQTILETVQEHPDLNAVRLGISGFSLOCHTAFELAKIYN
PRDLNIKALSVMAPIADGGILLKELYENFSKHGEGDIISVGKDFGFGPPPIIVCSGDVDL
LIRIODHVTANSLPTKPYJLHQQGIDDTLVSRTQCTLFKNTAPGRMTFISYPNTGHNLAT

APDLDMILDQIVSHFORT/L

197892 195430 AKEAPKEEVLSLSRSVVAKYLENAQIRKEIYVPNKLVNFVL

197874 199202

CPI_0154

998A-WO Transferase

TSEFCPMMLRGVHRIFKCFYDVVLVCAFVIALPKLLYKMLVYGKYKKSLAVRFGLKKPHV
PGEDFLWWFHGASVGEVRLLLEVLEKFCEEFFGWRCLVTSCTELGVQVASQVFIPMGATV
SILBLDSI IIKSVVAKLRPSLVVFSEGDCWLNFIEEAKRIGATTLVINGRISIDSSKRF
KFLKRLGKNYFSPVDGFLLODEVQKQRFLSLGIPEHKLQVTGNIKTYVAAQTALHLERET
WRIBLRLPTDSKLVILGSMHRSDAGKWLPVVQKLIKEGVSVLWVPRHVEKTKDVESSLR
LHEPYGLWSRGANFSVPVVVVDETGLLKQLYVAGDLAFVGGTFDPKIGGHNLLEPLQCE
WPLIFGPHITSOSELAQRLLLSGAGLCLDEIEPIIDTVSFLLNNQEVREAYVQKGKVFVK ETASFDRTWRALKSY I PLYKNS

CPD_0155 199697 199488 No pobust homolog present in Genebank/EMBL as of 11/7/98 NSESFGVPFLEKLKISLIPIEEMRHELFMKTHNSSSNGFSNQEKGIRTYFKSDLLGYEDL 199697 199488 YFLRENINPN

200147 CPn_0156 199770 No robust homolog present in Genebank/EMBL as of 11/7/98
LGKOKLLARHMHNIVVLSEEGGRSAFLGRTAFFPNKYPIACGGVGIPSTIGNLFTIWYC
FYFYRAATPQSDHPDGCGFILLERLKELGAGFFYCDLRESNTTGFTLFFEGSNKGVLKNH

CPn_0157 · 200753 200298
No robust homolog present in Genebank/EMBL as of 11/7/98
FSFVTYKEALMNIYOFSPGASFNWGASLMAQLNSYFCLGGETVTRIISLRPSGLILAKKE
KAVVSTAEKILKILSFILFPLULIALAIRYLLYNKFNKDLDRAVFFIPTEITKAEELIIA KNPALVKEAALTVSPLFYSLPKKYQLMKVETP

201463 200894 CPn 0158 CPT_0138 Z01463 Z01034
No robust homolog present in Genebank/EMBL as of 11/7/98
PPKITLSINIDLLLEDLDTDSIPMPKLYLSEDFDFAYYPESKAIIDTVAKLEKNNPGEEF
CLESKKILARYLLEQLFKLETGLNFPTSTIDGGRESFLIEFSHETKKPTVWAFIVFYYH SNGPKLEKDFKQAGCEVHNRLLNLGLKYRPQAGAQNDGRNGGPYGP IGFLIVWEENYGSV LKDHGFIKDN

CPn_0159 201811 201467
No robust homolog present in Genebank/EMBL as of 11/7/98
CCPOGETATRIFSMTPSGFSLATEEK/VOYSTAEKVIKILALIFFFIILIALAIRYFLHRK
FDRKCEVIPCDTPKELELILAANPOLVEKAAREVHPGFFALPTKYQSMYIQTSKG

CPL_0160 203794 202127
pfkA-Fructose-6-P Phosphotransterase
TVELLSLNKSYFEIGRLRYRPEILTLLETIRSKHIQETSSPPSPPPELQKHIPNLCRIPE
VSIYTEOETSSKPLKIGVLLSGOGAFGHNVVIGLFDALRVFNPKTRLFGFIKGPLGTR
GLYKDLDISVIYDYYNMOGFDMLSSSPEKIKTEEQKKNILLTVKQLKLDGLLIIGGNNSN GLYKOLDISVIYOYYNMGSFDMLSSSERIKTEEQKKRIKTIVKQLKLUGLLI IGGNISN TOTAMALBE FLANKCKTSVIGVRYT ILOJOLKKOW IETSGEFITSGRYTSBEIGNLAKOLS AKKYHIFIRLMOQOACYTTLECGLGTLENIAL ISELIATRKISLKOLSEOLAGLVRRY KSCKNYCTTVLI PEGLIEH FEDTRKLI GENTULLANGDISJERILGKESPETLKTFHLEFK DIANGLKARDGIGNYNSKIKTREELLANWYKKETERKIKTHEFKY FOXFARAFF SNEFCNYGIALGIGLISALEVROKTCYM IT INNIAUSYYTEWCXIXTPLYKMMHLENRCCTE TPVTRTDSVDPKSPAVQHULQQCDCCL/EDLYRFYCPLQYFCKEELTDQRPLTLLWENQTHSPPQALYSTSGKRSL

CPn_0161 204658 203798 (predicted acyltransferace family) HRISTSTRIKEEPEL POJVALERHEDETMESTETLENPTTFGILHTTEHVINPTYFTVILLIG LASDATOSKRSHVRLAQELTREGIAALRADILGIIGDGEGELMDFSSENYRQHIJEETEYT HSLIHITQERLAIFGSSEXSTLADGFLIFFNKIKALAAWALTISKELMAAFAQENAPEVI TM::QKGAITYAGMTLNPDFYTQFLKIDIVKELMPSARI FTEAFANQDKPITILTYPDVDHAFPFAESSALSDLTOI

YMQGEQDLLVSINHRTL TSGE

CPn_0162 205870 204803
No robust homolog present in Genebank/EMBL as of 11/7/98
EVYTUNIQSPFRIMKLYSISSDUTPWIFQLMSKVDSYLFLGGRIKVVSIVMQEPNLI
IGKVENVRISTIVKILKILSFLIFPLILIALALHYFLHAKYANHLLVSKILERAPGYVPI
MRSGDTASSHYKLTTLVPVSOKNLQAMGSNPLEVEAALRTTKPSFCVPAKYRQIIISSH
LIBECHELBAGAGHILD, TOWNTEVELBTMDEF ERADERVLYMYZHLFTTYTYINGVER
LIBECHELBAGAGHILD, TOWNTEVELBTMDEF ERADERVLYMYZHLFTTYTYINGVER
LIBECHELBAGAGHILD, TOWNTEVELBTMDEF ERADERVLYMYZHLFTTYINGVER
LIBECHELBAGAGHILD, TOWNTEVELBTMDEF ERADERVLYMYZHLFTTYINGVER
PLEZEMICHTUNGVEFFICPSGGPDNPNFQGFQGIRIYWEDSYQPNKEV 205870 204803

CPn_0163 205831 206394
No robust homolog present in Genebank/EMBL as of 11/7/98
FEKAIVYCIKCKGIIKCISIIHTPTPATELCTEGEIFFGPVDSAIQNDLERLLTVKKRFD
IREYLRAGGSLVTTYPKEGORLRSPEQLRVLDDLVQSYPNHLHAIELDCGAIPQDLIGA
TYIITFADFSTYILSLRSYQANSPSDDTWGIWFGSIDDPVQAVISFLKDHGFALPSTLAQ

CPO_0164 206444 206998
No robust homolog present in Genebank/EMBL as of 11/7/98
LCFKCIYIKIIFSFLKQLMTRSTIESSDSLCSRSFSQKLSVQTLKNLCESRLMKITSLVI
AFLTLIVGGALIALACGGVLSFPLGLILGSVLVLFFSIYLVSCCKFFTLKEMTMTCSVKS
KINIWFEKQRNKDIEKALENPDLFGENKRNVGNRSARNQLEMILHETDGIILKRYMKGAK

CPn_0165 206983 207592
No robust homolog present in Genebank/EMBL as of 11/7/98
NVLLFNNWVPKTIDHVDPESEIDIRKVVSCYKLIKECQPEFRSLISELLGVIRCGLRLLK
RSKYQEQARTVSDEDAPLFCLTRSYYQDGYLTPLRAGPRDLINHYIHLRRRENPKHFFSP
KHPCYYARLAFNESVCVYRELFDIERLTKMYVEGDYSKEQEKNLQAILSFVKTLDEGKDF LIEHKDTDLIGRGFTDVFCT

CPn_0166 207594 207962
No robust homolog present in Genebank/EMBL as of 11/7/98
NCLKQYNKSDSIMSESINRSIHLEASTPFFIKLTNLCESRLVKITSLVISLLALVGAGVT
LVVLFVAGILPLLPVLILEIILITVLVLLFCLVLEPYLIEKPSKIKELPKVDELSVVETD

CER_0167 208309 207977
Nå_Fobust homolog present in Genebank/EMBL as of 11/7/98
NLWSHFPRGFFMLPFCPTILLAKPFLINSENYGLERLAATVDSYFDLGOSQIVFLSKQDQG
ITWEELSAKDRKFKPGSWNCTLYTEDPILPAKNSFSNCSDIQMRTPISPIH

CPE_0168 208716 208417
Nd :Gobust homolog present in Genebank/EMBL as of 11/7/98
SYNULRRRENPEHFFNPGHPCYYARLAFNESVRIYRKLFNTAELKQMYGAGDYEQQNEDN
LKGSLSFVQILDEKDGFDDFLATHKDTTFIGRGGADIFCS

Cintin 0169 209537 208710

NO. FOUND THE PROPERTY OF THE PROPE

DSEEGC EARAGEGIETRTIFSSAKVNPEGLIEITRVTFSSLKRKOPSSLPIRVTCQLG

CPh 0171 212444 211149 CPNE-0171 212444 211149

*gual-GMP Synthase
IIKLOSARRHLNTIFILDFGSQYTYVLAKQVRKLFVYCEVLPWNYSVQCLKERAPLGIIL
SGGPHSVYENKAPHLDPETYKLGIPTLATCYGMQLMARDFGGTTSFGYGEFGYTPIHLYP
CELFKHIVDCESLDTEIRMSHRDHVTTIPEGFNVIASTSQCS:SGIENTKQRLVGLQFHP
EVSDSTPTCNKILETEVQEICSAPTLWNELYTQODLVSKIQDTVIEVFDEVAQSLDVQWL
AQGTIYSDVIESSRSGHASEVIKSHHNVGGLPKNLKLKLVZELRYLFKDEVRILGEALGL
SSYLLDRHPPFPGGCLTIRVIGEILPEVLAILRRADLIFIFELRKAKLVDKLSQAFALFLP
IKSVSVKGDCRSYGYTIALRAVESTDFMTGRWAYLPCDVLSSCSSRIINEIPEVSRVVYD
ISDKPPATIEWE

CPn_0172 213237 212440 cogenase (C00H-terminal region only)
APIGAAIGIGPLGISRAHHLVEAGANVLVIDTAHAHSKGVFQTVLEIKSQFPQISLVVGN LVTAEAAVSLAEIGVDAVKVGIGPGSICTTRIVSEVGYPOITAITNVAKALKNSAVTVIA DGRIRYSGDVVKALAAGADCVMLGSLLAGTDEAFGDIVSIDEKLFKRYRGMGSLGAMKQG SADRYFQTOCQKKLVPGGVEGLVAYKGSVHDVL/KQILGGIRSGMGYVGAETLKDLKTKAS

FVRITESCRAESHIHNIYKVQPTLNY 214041 2137/15 thr_off)

No robust homolog present in Genebank/EMBL as of 11/7/98

TIFDLIYKIDSYKHQQGFMDFSVFPDRFV#STSPSPIEDIDAKTLVSNCCHYCSRCLFIF

LULLSIIIOFSVYGTSGETASLVFGILSYIVLVLLIIECRNRECCRRIS

TO TODUST homolog present in Genebank/EMBL as of 11/7/98
KIFIMPYEKIVILSMIMTTISSPSSALMPELSLIEPPTLVSSTTOTLATTIPAOGRRS
TERKITULIFITISTENSPSSALMPELSLIEPPSTLVSSCOLTIVGLELIMGLYFM
ISSLOZ/SLVGLIQKELSQAEEREFEYIQETEALRGAPRAESPTES

2148% CPn 0175 215275 THE COURT HOMOLOG PRESENT IN GENEBANK/EMBL AS OF 11/7/98 LILACE/FILERPEMEQUNCY/ODTTVLYALNSEDERLSDETHREGKOSPLEAENALGE FIFELLTHEEPPLEEVAIPILEOVALPHEKFYLSFIDRDGOGVHYEVLDGVFLKTVAACIIENS FLTDSMSPELLSEVKEALK

CPn_0176 215271 216518
CT153 hypothetical protein
NDDDPMDESDGEASKDSAFSAS/SYEFVKSSTRESKNTVTHSTASRTLYILRODCSYDP
RALKVDDEFRYWEKKLDANNPUSLNAFVKEVGTHYVASVTYCGIGFQVLKMSYLQVEEL
EKEKISISVAASSLLKSKTONATEKCYSSYGSESSACTVFLGGTVLPDLQQDKLDFKDW
SESIPNEPIPLAISVSSITDLIFELFFSEDAQVLSQKKSALDQVILNYLESHKPKEEGP
PLAYITUSGEHGGGTGTTLA AFAPPTASPTYLTYVACTIFVLFFTLFFTCAQGELSFLBF
BITTEROOMICAUTTUS LAAFTYBLIVEN BETERDFTSTLFTCAGGELSFLBF
LIGHTYGDLFTRSGDEIBLKHNTSGKTLAYTGMJLSTCKNLTCTTGTSDVFFTITY

CPn_0177 217513 216608

No robust homolog present in Genebank/EMBL as of 11/7/98

DKREQTKSKFIFLISEESMKOPMSLIFSSVCLGLGLGSLSSCMOKPSWNYHNTSTSEEFF
VHGNKSVSOLPHYPSAFKTTOIFSEENDDYVVAKTDEESRKIWREIHKNLKIKGSYIPI
STYGSLMHPKSAAL/TLKTYRPHPIWINGYERSFNIDTGKYLKNGSRRRTSHDOPKNRAVL
NLIKSSGRCNAILLEMTEEDFVIARREGVYSLYPVEVCSYPQGNPFVIAYAWIADESA
CSKEVLPVKGYYSLWWESVSSSDSLNAFGDSFAEDYLRSTFLANGTSILCVHESYKKVPP

CPn_0178 218052 217789
No robust homolog present in Genebank/EMBL as of 11/7/98
VKEYLDFLYORNVERDPOTKRHCTVSOKFGGESIDAKTTTGQLFHIAGKTEPGHGKLCLG
ESILKQLYALGIITGYENREREVWVYLD

CPn_01/9 218550 218056
NO robust homolog present in Genebank/EMBL as of 11/7/98
PKIWDTHFETRIEATSVPKFNRTLRKSFHKSGRSSRPSKACVANFFNFTLQAGRSGIIPG
KKAJLLNVNDAKTPNYSCIFESIGFFNEQDLEAQHNQQAALVRKILKVVPHHFLKGLIAK
LPPSLKKDRKFMSSLIFTKLSYALDLSAPMHLEGKPNLSYEEKLD

CPn_0180 218963 218355
To robust homolog present in Genebank/EMBL as of 11/7/98
TSLHKILDCKYKPVFIQNTVASETYPSQILHAQREVRDAYFNQADCHPARANQILEAKKI
CLLDVYHTNHYSVFTCVDNYPNLBFTFVSSKNNEMNGLSNPLDNVLVEAKVRRTHARNL
LAACKIENIEVPRVVGLDLRSGILISKLELKQPOFOSLTEDFVNHSTNQEEARVHQKHVL LISLILLCKQAVLESFQEKKRSS

CPn_0181 219175 218777
No robust homolog present in Genebank/EMBL as of 11/7/98
VHELFKIDGVYYFFKKFMKLFYNNYSLNSHHEKPSSLEKAVOALDSYFYWGGDTTDVLAR
DDISREIYCVRRLYIFFWIVSISOSLSRIPWRLKRILLRYCTLRGKYVMPILIKRIAILL

220704 219334 CPn_0182 220704 219334
accc-Biot in Carboxylase
RCIMKKVLIANRGEIAVRIIRACHDLGLSTVAVYSLADQEALHVLLADEAICIGEPQAAK
SYLKISNILAACEITGADAVHPGYGFLSENANFASICESCGLTFIGPSSESIAMMGDKIA
AKSLAKKIKCPVIPGSEGIIEDESEGLKIAEKIGFPIVIKAVAGGGGRGIRIVKEKDEF
RAFSAARAEAEAGFNNPMYYLEKFIENPRHLEIQVIGDTHGMYVHLGERDCTIQRRQKL
IEETPSPILNAEIRVKVGKVAVDLARSAGYFSVGTVEFLLDKDKKFYFMEMTRIOVEHT
ITEEVTGIDLVKEQIHVAMGNKLEMKQKNIEFSGHIIOCRINAEDPTNNFSPSPGRLDY
LPPAGPSIRVDGACYSGYAIPPYYDSMIAKVIAKGKNREEAIAIMKRALKEFHIGGVQST
IPFHQFMLDNPKFLESNYDINYIDNLLAQGNSFFKEF

221207 220695 decab-Biotin Carboxyl Carrier Protein
RRICMDLKQIEKLMIAMGRNCMKRFAIKREGLELELERDTREGNRQEPVFYDSRLFSGFS
QERPIPTDPKKDTIKETTTENSETSTTTSSGDFISSPLVGTFYGSPAPDSPSFVKPGDIV SEDTIVCIVEAMKVMNEVKAGMSGRVLEVLITNGDPVQFGSKLFRIAKDAS

CPn_0184 221814 221221
efp-Elongation Factor P
OWKIKFCCCEEKIMVLSSQLSVGMFISTKDGLYKVTSVSKVAGPKGESFIKVALQAADSD
VYLERNFKATGEVKEAGFETRTLEFLYLYLEDESYLFLDLGNYEKLFIPOEIMKDNFLFLKA
GVTVSAMVYDNVVFSVELPHFLELMVSKTDFPGDSLSLSGGVKKALLETGIEVMVPPFVE IGDVIKIDTRTCEYIORV

222457 221765 rpe/araD-Ribulose-P Epimerase AEVKKQESVLVGPSIMGADLTCLGVEAKKLEQAGSDFIHIDIMDGHFVPNLTFGPGIIAA INRSTOLFLEVHAMIYNPFEFIESFVRSGADRIIVHFEASEDIKELLSYIKKCGVQAGLA FSPDTSIEFLPSFLPFCDVVVLMSVYPGFTGQSFLPNTIEKIAFARHAIKTLGLKDSCLI EVDGGIDQQSAPLCRDAGADILVTASYLFEADSLAMEDKILLLPGENYGVK

222878 224068 *similarity to Cps IncA
PIKDKILMSSPVNNTPSAPNIPIPAPTTPGIPTTKPRSSFIEKVIIVAKYILFAIAATSG
ALGTILGLSGALTPGIGIALLVIFFVSMVLLGLILKDSISGGEERRLREEVSRFTSENQR ALGTITALISGALTPOIGIALUTIFFYSMVILLELIKUSISGALERIKREEVSKI ISSAVA LTVITTTLETEVRULKAAKDOLTLEIEAFRNENGILKTTAEDLEEQVSKLSGOLEALERI NQLIQANAGDAQEISSELKKLISGMDSKVVEQINTSIQALKVLLQQEMVQEAQTHVKAMQ EQIQALQAEILGMHNQSTALQKSVENLLVQDQALTRVVGELLESENKLSQACSALRQEIE KLAQHETSLQQRIDAMLAQEQNLAEQVTALEKMKQEAQKAESEFIACVRDRTFGRRETPP PTTPVVEGDESQEEDEGGTPPVSQPSSPVDRATGDGQ

CPn_0187 234218 225045
predicted methylase
vpLTYTRTLPHHSKPLSRRKKNSSHKEETCWDGIASSYNKIVODKGHYYHRETILPOLLP
SLTLOSKSSVLDICCGOGFLEPALPKECRYLGIDIGSPLIALAKKMRGVNSHOFKVADLS
KRLEFVEPTLFSHAVAILSLONMEPPISA INNTATLLEPLGOFFIVLNIHGERIPRASSW
HYDENKKAISRHIDRYLGSHKIPI IMANIMOKODOPTLOPHFPLCTWPKELSGHGFLVSGL
EEWTCSKTCTGKRAKAENLERKEFPLFIMICKIKIK
CDn_0188 25990 336400 224218 225045

CTT_12 TAN
CTT 32 hypothetical protein
RTLBC HMFRKLEPFCKKKTG/RORLRNNGLLOAT [O]: [KV], DINIEACKEACVLEYYGLL
PCVPTLVFFLRGQHLFTNLWWEWLITKFPDYKKFT/ATVEAAYHATEEN IGLVLVGSF EVERMACI LINGUICE SUNKTERTOWTE GUPPLICITY PT OF THE COMMENTAL STUTIO IMPTOYAKLESUSHSMTALYETOREVEYILLI YLADEC (YAFLEPVA TOKTSAL ISTUTTO SVIMI VEOKAFESULVISTENVISETYGALVALECIFILLU YTYMI YLEFXIARTET FUNKOCT

81

FIFLGDKILDSCYLQLITSTYILALTTRGFNEGLSPLT EKEGFLFFYNNGYOPVFNFSELTIKDIADKLLHREIFK SKNKENLTLGEIARRIK

OSKVPIGEVSQCLDVL GITFIENSFQNIFNQA

CPn_0189 226391 229825
CT131 homolog (Possible Transmembrane Protein)
ANOMKRESMIKIIGICIASSIVIGELIFLPOLLSTESRKYLVFSLIHKESGLSCSAEELK
ISWPGROTARK IKLTGEAKDEVFSAEKFELDGSLERLIYKKRKGITLGGWSLKINEPAS
HARDWEILDBEALD BEACH DER EINE STEMEN STEMENSTEREN FOR STELLIYKKRESTILSGWSLKINEPAS
HARDWEILDBEALD BEACH DER EINE STEMENSTEREN FOR STELLIYEKRESTILSGEREN FOR VAFFELBLADE WHILLEWESTENDE FOR FOR THE NOOK ISMTASGEGNOIOMKLOCHINKSTFYIVEGSSSFIELKRELASALCNOIIPLSTE
TISKOIHATVSYAKT PLDITKWKH EITGOOGLPEVAIHFKDDRIALGLARTKLGIKKTE
KFSDIRYSSSTVLGGASPSHLNGLISIDNKKHLTKFRLOOAGLPHTURAIFROPFVINV
PLDVAYYSLNIEGTYKNAHLEADAILDNELKLSCSMSGTWKNFLFKGGOTYHFNKMOE
FLSPHFSVARAFFSKAGAITDTNIFFPKFSGK ITARENELLIHAKFGSPNEP IKPETTSI
LIHOGFCSLPLSLVSNHLAPFHLKKLTFSFHTDGGKFVTKGNLOALIENEDY PDLANTRI
LIPDLLISLDESSTSPSKDLKIOGSGEIFSLPLDSITKTYGKOVLSPYFGSSGDLNFV
VMYNFKDONKLTLLSNFKSEALLGELKLVMDFSMKLSSCTOCTLOMEVSPERVASFFKNA
SCSPTCLLHRTANVRLDISKLSCPETKGLSCLTLLAAGGLEGSLEATPLIFYDNVSKT
FILNDFKGSLRANNLDAXIETYDLKGSCLAPRODSKTLAEFSLEGOVDHLFSPESREFKOT
ANWIHIPSSFIAGIIPMSPGLKAQISSLAGPRINVSIKNAFRGEGPVDIMVDSBHJOKA
PLILNEKSILLRENLTAHLSINEDVNKAFLQEFNPLLAGGAVSOYPVTLEIDKONFYLP
IRPYSFEEFRIGSATLDWKKISIANTGTMYALFQFLDITDQKGFVESWFTPIFFSVQKGS
LICKKLDALIDRRIHLALWGKTDIAHDRLFMTLGIDPEVIKKYFHNTSLKTKNFFLIKIR
GSISSPEVDWSSAYARIALLKSYSLGNPFSSLADKLFSSLGDSTPPPTVHPFPWEKSNFD

CPn_0190 229901 231274
No robust homolog present in Genebank/EMBL as of 11/7/98
LLGIKLMRKRHSFDSTSTKKEAVSKAIQKIIKIMETTDPSLAVETPNAEIESILGEIKEI
KQKLSKQAEDLGLLEKYCSQETLSNLENTNASLKLSIGSVIEELASLKQLVESIEESLG
QODQLIGSVLIEISOKFLSSIGETLSGNLDMNONVIQGLLIKENPEKSEAASVGYVQTLL
EPLSKRIGETHKKVATHDVNISSLQFHMMSVAGGRFRGHIDMNSYRVLGLGEPKNGEDAV
SKDYLERYVSSQLTIDKVEDKPITKPNKGKLLYSQGTSPKLESPLPLGLLTSGISGFMX
SASKSNDGSFPFSALRHKETESDTDCFQITSTTLSGNQAGTYTWSLSLKVLVPSIFQIEK
PEVQLSLVYSYEDWLPIDNIFMSQPRITPLALLGGTMLAGQKYDILELAAHQTNQTLMI
SPNCSRFSLQLKQTNQFENSPVDFYIVHAAHSCHWSGF

CPD_0191 232039 231314
ging-ABC Amino Acid Transporter ATPase
OfffpryeLdyOKREGYMTIRVENLAYSVNKKKILLGYTFSLERGHITLFVGKSGSGKTMI
LRALAGLVOPTOGDIWIEGEAPALVFOOPELFSHMTULNCTHPOIHIKGRSTEEAREKA
FELLHLLDIEEVAKNY PDOLSGGOKORVAIVASLCHOMHTLLFDEPTSALDPFATASFRH
LLETLROQELTVGLTTHDMOFVHSCLDRIYLIDQGTVAGVYDKRDGELDSGHPLSKYIHS
AQ
ii

232643 231984
ghh-abc amino acid Transporter Permease
evgyphmlatarillegcgyticvsGificgsIlGiltgtvtsLyfpskltkllansyv
tvTrgtplfigiliiyfglp5ulpieftplvagiialsknsaaylaenirgginsLsig
wesamvlgykkygifyiiypgvfknilpsltnefvslikessilmvvgvpeltkvtkul
vsrinpmemylicaglyflmttsfscisrlsekrrsynn

CPR 0193 233144 232686
*argR-Arginine Repressor
KLHGVFMKKKVTIDEALKEILRLEGAATQEELCAKLLAQGFATTQSSVSRWLRKIQAVYV
AGERGARYSLPSSTEKTTTRHLVLSIRNASLIVIRTVPGSASWIAALLDQGLKDEIJGT
LAGDDTIFVTPIDEGRLPLLMVSIANLLQVFLD

CR1L0194 233162 234241
gcp-0-Sialoglycoprotein Endopeptidase
EVERTIKGNYFFSWFFMLTLGLESSCDETACAIVNEDKQILANIIASQDIHASYGGVVPE
LASRAHLHFPOVINKALQOANLIEDMDLIAVTDYPGLIGSLSVGVHFGKGJAIGAKKS,
LLGTMHVEAHLYAAYMAAQNVOFPALGLVVSGAHTAAFFIENPTSYKLIGKTRODAIGET
FDROGRELGLPYPAGPLIEKLALEGSEDSYPFSPAKVPNYDFSFSGLKTAVLYAIKGNNS
SRESPAPETSLEKQRDIAASFOKAACTTIAOKLPTIIKEFSCRSILIGGGJAINEYFRSA
IQTÄCNLPVYFPPAKLCSDNAAMIAGLGGENFQKNSSIPEIRICARYQWJSVSPFSLASP

CPT_0195 234172 235785

OPPA-Oligopeptide Binding Protein
YSGNSYMRKISVGICITILLSLSVVLOGCKESSHSSTSRGELAINIXDEPRSLDPRQVRL
LSEISLVKHIYEGLVQENNLSGNIEPALAEDVSLSSDGLTYTFKLKSAFWSNGDPLTAED
FIESWKQAVATQEVSGIYAFALNPIKNVRKIQEGHLSIDHFGYHSFLSVVLTLESPTSH
FLKLLALPVFFPVHKSQRTLQSKSLPIASGAFYPKNIKQKOWIKLSKNPHYYNQSOVETK
TITIHFIPDANTAAKLFNQGKLNWQGPPWGERIPQETLSNLQSKGHLHSFDVAGTSWLTF
NIKKFPLNNMKLREALASALDKEALVSTIFLGRAKTADHLLENNIHSYPENGKQEMAQRQ
AYAKKLFKEALEELQITAKDLEHLNLIFFVSSSASSLLVQL/REGWKESLGFAIPIVGKE
FALLQADLSSONFSLATGGWFADFADPMAFTLIFAYPSGVVPFYAINHKDFLEILQNIEGE
QDHQKRSELVSQASLYLETFHIIEPIYHDAFQFAMNKKLSNLGVSPTGVVDFRYAKEN

CPn_0196 235906 237519

oppA-01igopeptide Binding Protein
KLKSYSKERSFMLRFFAVFISTLWLITSGCSPSOSSKGIFVVNMKEMPRSLDPGKTRLIA
DOTLMRHLYBGLVEENSONGEIKPALAESYTISEOOTRYTFKIKNILMSNGDPLTAQDFV
SSWKEILKEDASSVYLYAFLPIKNARAIFDDTESPENLGVRALDKRHLEIGLETPCAHFL
HELTLDIFFPVHETLRNYSTSFEEMPITCGAFRPYSLEKGRLHLEKNPMYHNKSRVKLH
KITVOPISNANTAAILFHKKKLDWGOPPWGEPIPESIASLHODDOLFSLPGASTTWLLF
NIGKKPWNNAKLRKALSLAIDKDMLTKVVYOG/AEPTDHILHPRLYPOTYPERKRONERI
LEAQQIFEEALDELDMTREDLEKETLTFSTF9FSYGRICOMLREGWKKVLKFTIPIVGGE
FETIONNELGDNYSLTVNOWTAAFIDPMSYLMIFANPGGISPYHLODSHFOTLLIKITGE
HKKHLRNGLIIFALDYLEHCHILEPLCHEN RIALHKNIKNFNLFVRRTSDFRFIEKL

CPT_01177 2.375.12 2/8882
opps of igopopt ide trinding possible filderoperties for the knyrrestableroperties terminerby the led to the construction of the constructio

ELIPTYYLSYDYLTODFIET

CPN_0198 237 24076

ODDA-QLIGODEDLIDE BINDING
CPT_0199 2/1018 241983
oppB-01igopeptide fermease
KKLIGLSLVFSYIKNRILFNLLSLWIVUTLTFLVMKTIRGDFFNDEGGNVLSEEVLQTLK
KRYGLDKRLYQOYTOYLMSIAKLDFGNSLVYKDRKVTNIISTAFPISAILGLQSLFLSIG
GGIALGTIAALKKKORSYILGASILQISIPAFFITALLQYVFAVKIPLLPIACWGSFTH
TILPTLALAVTPMAFIOLTYSSVSAAINKDVYLLAYAKGLSFLKVVIKHLIPYAIFPTI
SYSAFLTITVITGTFAIENIFCIPGLGKWFICSIKQRDYPVALGLSVFYGTLFMLSSLLS
DLIQSIIDPQIRYYGKEKKKK

CPn_0200 241996 242868
oppC-0ligopeptide Permease
ekkkhkkimenkssapsrsimksijonkmivigittliilmigallipwfyqdyeqtsikb
ilvspcsrfpfgtdtlgrcmfartirgirlisiliatiatlidvcvgilwatvaisggkki
bfilmarttekifsipripiilillivifhhgilplilamtittgwipisrijygofellikm
pfvlsakaphastfhiikkhilpwilapiistliffipraivteafisfiglgippqas
lgtlvkbginaidyypwlfffpslimialsisfnligegakticleegshg

CPn_0201 242810 243715
oppD_01igopeptide Transport ATPase
ASISSARGLKHYVSKRDLMGNYLINIKDLTITSTNFKRTLIENLSLQLKENRNLALVGES
GSGKTTITKAILGFLPENCLIKTGSILFEDIDITKLSPKELHKIRGGKIATILONAMGSL
TPSWRIGMGIIETLROHHKMNKERAYNKAMQLLTDVCIPNPKYSFSQYPFELSGGMRQRV
VYATALASOPKLILADEPITALDSNSQAQVLRILRNIQOOKQATILLVTHNLSLVKELCN
PICIIKDGKLIETGTVEEIFLSPKHPYTLKLLNAVSKIPIKKTSSPILKNKFQPLMSMOG

CPn_0202 243682 244500
oppF-0ligopeptide Transport ATPase
vPTSNEYARMFMTTLLSIKDLSLTIRGKKILNHINLINLIKGSYLTIVGPSGSGKSSLALT
ILDLLKPTTGTITFHMDPKIPRARKVQVIWQDIDSSLNPCMSIKGIISEPLNIIGTYSKA
EQNKEIYNVLDLVNLPKSVLHLKPYKLSGGQKQRIAIAKALVSKPELLICDEPLSSLDTL
NQSLILDLFQTIKKEYQNTLLFITHDMSAAYYIADTIAVMDQGSLVEHACREKIFSTPKH
TTTQDLLDAIPIFSLISTEMEPSEEYELQVASK

CPn_0203 244966 245802
No robust homolog present in Genebank/EMBL as of 11/7/98
IVPLPQKNNKETSCMNTYTFSPTLQKSFSLFLLEKLDSYFFFGGTRTQILVITPTNIRLA
AKKRGCKVSTIEKIIKILSFILLPLVIIAFILRYFLHKKFDKOFLCIPKVISNEDEALLG
SRPQAVEKAVREISPAFFSIPRKYQLIRIDTPKDDAPSILFPIGIEIILKDLCIDTLKQS
NLFLKREMDFLGHPEEKALFDSICSIEKDQEMMSLESKKLLITHFLKYLFVSGIEQLNPG
FNPENGRGYFSEISTAKIHFHQHGRYGPIRSSGPIMKEI

CPn_0204 245691 246002
No robust homolog present in Genebank/EMBL as of 11/7/98
PREWAWVFFRNKYSKDPFSSARSIWANPF7GTHHEONIKIKGMGYQIFTRLKKLGISFSS
YNSINPNYFFDEGCFVYWESOFKSALQDHGILQKQTETFYRNT

CPn_0205 246073 246327 No robust homolog present in Genebank/EMBL as of 11/7/98 IEDSIKGYGSASAFRNPPOLLLKFFLVCEELCILTVATHRALLETPLALSFFKELKTKYV YRAKDILQLHNYKGFTILNTSPLCS

CPn_0206 246346 247161
CT203 hypothetical protein
IVDRRSPACYDS INSDAIGVSLLMDISHILEDLAYDEGILPREAIEAAIVKOMQITPYLL
HILHDATORVEEIVNDGSYCOHLYAMYLLAQFRESRALPLIIKLFAFEDDTPHAIAGDVL
TEDLPRILASVCNDDSLIKELIETPKINPYVKAAAISGLVTLVGAGKIPRDKVIRYFAEL
LNYRLEKOPSFAWDNLIAGICTLYPGELFYPISKAFDGGLVDTSFISMEDVENIIHEETV
ESCIHTLCSSTELINDTLEEMEKMLEDFPIEP

CPn_0207 247209 248617
ybhf/soditi-0xoglutatate/Malate Translocator
vnkkkrflsllfltavlljiwfsphpasinsnawQlfaiftttimgiifqpvpmgaiaii
gistlltotltleoglsgfhnpiawLvflsfsiakgiiktglgeriayffvsalgkspl
glsvicvitofflafaifsvtapaggiltpvvtslsdsfgssaekgtqDligsflikvay
gsvitsamfltamagnelvaalaghvovslswvLwakaaiipcleslflmpiiltyklyp
pkitocefairsakLrlemgplkkeektilmiffllv/Lwftgdbligisattaaligls
llitnildmokdvianttawetfiwfgalimmasflnqlofiplvodsaalvsglswl
ligfpllfliyfyshylfasntahigam/piflavsislstnpiffaaltlafasnlfgglt
hygsgpaplyfgshlvtvoewwpsgfalsivniviwigigslwwkalgli

CPI_0208 24935 250602
pfkA-Fructose-6-P Phophotroinsferase
svavilmipplyvdlottissyspplpkepobassliavpdtshskpvvpgvktlppqtyh
LPYLKFVQGENVViitplkvgvmfskrppapxchnvioglpnslkdphpdsslvgfvnn3dg
Lthiks iditeefliskfsnskopen cytorkkivppbakbaclktabaldloduviggd
chthiks iditeefliskfsnskopen cytorkkivppbakbaclktabaldloduviggd
chthat labeyfakrristiydpyrtidbioloitplldtepgdptatfyssiisnisk
dalockahyhfiklmiskabhalalealaythehialaldeelaeknlplktiihkicsvia
dbamekyygvilifeatisfipbiinlithiblegebybkibrldeborllksppapi
legitindbalkonyyksisvbyllihilminniboyffpippphatshefloyegrsolptk
pump gysloygabilaknbespytetteblagpphymiralfonfyrdkogadotto
pkipkyllvdigstabrkspkypymalebospppidebyfettebmhodppplttllnhn
pkophosicetepptypi

CFm_0209 251000 251000 Account for General REPRESENTATION OF CONTRACT CONTRACTOR OF THE CONTRACTOR OF

HILDMITYK ILKYLSEPRSLLRTTSLWYRP

CPn_0210 252459 251440
No robust homolog present in Genebank/EMBL as of 11/7/78
YQKLWEREREYFKTIREKEHATISTMLVELEALKREFAHLKDOKPTSDOEITSLYQCLDH
LEFVLLGLGGDKFLKATEDEDVLFESQKAIDAWNALLTKARDVLGLGDIGAIYQTIEFLG
AYLGKYMRRAFCIAJEIHFLKTAIRDLNAYYLLDFRWPLCKIEEFVDWONDCVEIAKRKL
CTFEKETKELNESLLREEHAMEKCSIODLORKLGDIIIELDDVSLFGFKTPSQEEYQKD
TAGGERIASLAFETTALEFTTATOPORGAAZKEFIREKTGLAEGSZTIPQTKELEFAI
ARCHLEBERGLYMPKYEJAAEHGLTSWHEEBSTTTIIGSSTEETTATE

CPn_0211 252765 252463
No robust homolog present in Genebank/EMBL as of 11/7/98
ECVMSYPDISNVQASSIQSALLHKTSDOIQOKRCFKQSTFVILAVSLVIIGSLFLLAGVA
ILTVFSHGVLSLVFGVLGIVLGLLLLAGGVGLLVEEAKSLL

CPn_0212 254081 252888

No robust homolog present in Genebank/EMBL as of 11/7/98
ELSYGVVVSIYSEILSFSELTSCKHSLFPFGFIETASIRIHHYFNVVIVCLIILGTLFVC
LGMYFLGVFSTYLLGMSSMILGLLLISIGLALLKFKERYGLEFKELFGVEGGFDKKLPSE
IIOMODQIADLARELDLEQKKDTLIRGFSARLDVLEGSKTEKKQILKIGVPRNLSEIQER
AQEONSILEQCKEALLFRRKSAQEIFKKLYDRKAAFWRSYREDLMCYSEIHVSKKALSNL
YIGDVFEGTAPHFLMEAYAMCRTAKNLRNYVKVCVEDMRVNEEKKRAKQLSVSELLCCCT
EIETDLENETNLFTSDSEDVLEEYQIHCIRVTMLHALWAIYNDEVVSRKPIDTLDRVRAR
MAVEDCIETFEELOMCVVHTKTLELEIAQLYVDILLEA

CPn_0213 254345 254190
No robust homolog present in Genebank/EMBL as of 11/7/98
ILVVFSRVIFSNTNQIGIPRLELILPLWKKENDPFCFLFSRVEGTFIILNIK

CPn_0214 255768 254446
No robust homolog present in Genebank/EMBL as of 11/7/98
FLGLKEDYERPTYCHI PPAPHPORVDSKGCIASHVSTVVVVALFILGIFFLSGSLAFLVH
TSCGVLLGAALPILCIGLVLLAVALIVFLCHKHKTRQDLDYYDQDLDSLVIHKKEIPNDI
SELRVIFEKLQNLFOFHTKDFSDLSQELQGKFINCMERKILTLEDEVITKFLIVRDRFLETR
RNFTTFGGVVKGIQSHIFDLHEEKSLVILEIYRLRRDLQVLLMFFLLPPGILKVDYDEIE
AIKGLFIRLTSRLDKLDVKAQERKKFINEMSREFKEVEKAFDIVDRATKKLMDRAKKESP
ARLFMGRTESLLEMKKNEEALKNQGLDFENLSHPELFSPYQQLLILMYLNSEIVLHHYEF
LISGTVTSGLITLEECENRWRAASTGLNALLVRKLQFRGAIKSAYFEKLTEIEKELRSLQD
VIKSLELELIHKIKDIVTEET

CPT_0215 257039 255759
No.Tobust homolog present in Genebank/EMBL as of 11/7/98
LT\$\$KKQVMSSATARDCFFSPSPQPSSTLGVMPPKYKSLILSVSLIVLGVLLLCVGMLLL
VM&FFSFSVLTVGLGGAGVFLGSLLLILGLIFFVSYHRKLSEATRSLEQKITLEYQPWAD
LAYELIEVQEWSNFLLIDEWEDFKEVVAQHKSQFATFEGDLLLFGREVEKYETIMKELDGR
DV&BLTELKNIWGPLEFLRKKGDRLQCEIDKLRKEVMKVGKSGLKLACELTKFKSALKDV
KL99GEYTROKRKYVEKLEVPFEGYRRELLEVLKTRLSVEKEQIFFEVVSAFERLASLHR
TV*SEEELQEALDKAKAELLDIQVRKSVVEDLSCEPTLIQVHLLRLYEVQCRIVEQFLTQ
TFSSEQEKVLEEYEALKARIRKTLRVKLDQVRANVAFVASTTDLLSESESLDGNDSVFED
AMDDFLD

CPT D216 257623 257174
NOT FIND 16 257623 257174
NOT FIND 16 257623 257174
NOT FIND 16 257623 257174
NKARTMNPVFFRIQUEFIPEDTSLRINSYIVACGLLILGVVLSILSVICLDIGLVGLSA
GAAFTLGLGCLIFALFLFSFSLILLLSQEKRVPDVLSLYLEKEVPQYETPLYKEDLESER
DWSAISFRIGIIEEKLRIAEKFRYSDSVFV

CPn 0217 257881 258579 ypdP

ypdb P PKECKLKGFLSVNELIFGFQTFSVVVLGVFFASRGKAWLTGWLSLLSSIMNVFVLKOCHL WGFEVTSADVYVIGLLTCLMYAREHYEKNDINDAMLCSWVISIAFLVLTQLHLFLIFSPN DS\$QEHFLALFSSTPRIVVASLVTLIFVQIVDIKLFTFLQRVFSKKYFAMRSTISILFSQ LIDTIIFSFLGLYGLVSNLCDVMIFAMLVKGIVITLAIPTLTVTKAVLDRRSS

CPn D218 259064 258582
Nolfbust homolog present in Genebank/EMBL as of 11/7/98
IFLSKKVFFESYEDFANVASSWPKSLRALVQGRYFVDSELKETPYRIHDFKXPPIHHRLY
RSLPIISTIGGIIRLIEAHSGPIHPRDKMKYRFEVLQAVIEILGLGVLILVFDIIGCFLA
FLVAIILSLLLYCNSTFTCVQNLSFTERMLEGIGEAVNFLA

CPn_0219 259348 260472

Cgt - Queuine tRNA Ribosyl Transferase
GSSLALKFHLIHOSKKSQARVGQIETSHGVIDTPAFVPVATHGALKGVIDHSDIPLLFCN
TYHLLLHPGPEAVAKLGCLHOFMGRQAPIITDSGGFQIFSLAYGSVJÆEIKSGGKKKGMS
SLVKITDEGAMFKSYROGRKLFLSPELSVQAQKDLGADIIIPLDEL PFHTDDGYFLTSC
SRTYVMEKRSLEYHRKDPRHQSMYGVIHGGLDPEQPRIGVRFVEDLFPDGSAIGGSLGRN
LQEMSEVVKITTSFLSKERPVHLLGIGDLPSIYAMVGFGIDSFDGSYFPKAARHGLILSK
AGPIKIGQOKYSQDSSTIDPSCSCLTCLSGISRAYLPHLFKVREPNAAIWASIHNLHHMQ
QVMKEIREAILKDEI

CPn_0220 260660 261238
No robust homolog present in Genebank/EMBL as of 11/7/98
FYSFLKKKGIFYMSKESIRSYSEISTPTPIFRETPSKEGVAYKLOLRSPAKDCILRNRVS
LKGALLRSIPFYGSFLGAKRIHSAMSAKDAPCTTRV/HYLVGGLELLGLGVVVLACKVLA
TALKFLFSKASSKIKQMKWREKARNLAAKDTVQSIKEFOSVDLTSCFTRCFRLRNRVVEE
GASENGTVREIIV

CPn_0221 261621 262064
No tobust homolog present in Geneburk/EMBL us of 11/7/98
DALRYKYEIGIGMVNRYKSGAFFACHYTYDODKLYBYCYKRNLRGLAPVENEVCLFEENNL
LEDVMASTFIMGHIGLIGREHISVMCTODFKDSKI/ITFHTALGILETLUIGIIVLLIKIT
TTILLILETPOLLCYFMYGAYSDFHPI

CTH_00220 262474 20284/ - weak dimitatity to Bacteriophage CHP1 (Ott4) GEKELKTWERLENAFELTOPEEYRHRWYMMTVLKCEPCRTOHAKVWSYRCVHEASLYE KHRFETTHYDDKHLPQYGGEVKLHLOLFLKELRKMIGPHKIRYFEXDAYGTKLORPHYHL LLS

CPn_0223 262950 √63333

No robust homolog p in Genebank/EMBL as of 11/7/98 immLigryschopereark, klapvronlegutneisetvsetsssikbsvirsleidssickark/itustnopedetqexiwhtifqaleticlgilillekiifvilheifHLVIGFCK

CPn_0224 263402 263674
No robust homolog present in Genebank/EMBL as of 11/7/98
YTFKNPKKNKKMKPNSIIFLENTKHYPDIFREGFVRDRHGLMEASDWLLSTEITIIRSIL

CPN_0225 263946 264541
No robust homolog present in Genebank/EMBL as of 11/7/98
NSFTIKFLLMTKNAINSOTTPOPNLTDAEPIASRAQCKSIAVIISLFALGMLLLCLGII
LISIPIPGLAAQVALGLGIYSLILGIALANIGFLCLLRCKQVPQPDTLPSESSKQPSE
GSTPTALPMQAGEFLEKVQVSATPILLPKNKDEELSAKVMKEGAEAASSIKQAVLESTEK
LIDARKQEESRREARKKIVAEEAEASRKRIQQQMAADQEALRKRKEEVAKRK

CPn_0226 / 264545 264967
No robust homolog present in Genebank/EMBL as of 11/7/98
AIFNRKRMPYYANTLEFIQCTOSLCPLFKYGFVRHHYKGQLEIEDASHDWDFLEPPSTWK
RTLLAAIPILGSVI&LGRLFSIWSIREPQDSQEYKSIFWHTLCAVLEILGLGIVALILKI
LATFIMAMPGLKRYATFLFYS

CPn_0227 265467 265009
dsbB-Disulfide bond Oxidoreductase
kERFNIFVSCYLLKEIMMINFIRSYALYFAWAISCAGTLISIFYSYILNVEPCILCYYOR
ICLFPLTVILGISAYREDSSIKLYILPOAVLGLGISIYOVFLQEIPGMQLDICGRVSCST
KIFLFSYVIAPMASVVAFGAIVCLLVLTKKYRG

CPn_0229 266242 265412
dsb5-Disulfide Bond Chaperone
VKDRADFINLKEKFSCSILKKENAFEFYYFCSIKOLTNSSLRGPLNKKILVLCTAMFFIV
CFGFLJHKKHTILPPKAHIPTNAKHFPTIGNPYAPINITVFEEPSCSACAEFTTEVFPLL
KKHYIDTGEISFTLIPVCFIRGSKPAAQALLCIYHHDPRQADIDAYMEYFHRILTYPKEE
GSHWATPEVLTKLAEGLKINSGRSVNPKGLEQCIASGQYNEQIKKNNLYGSQVLGGQLAT
PTAYVGDYLIEDPTFHEIERAIQHIRQLQAVEGDHDD

CPM_0229 266163 267560

CT178 hypothetical protein

#SKAFSFLRIEQENFSFKFKKSALSFTYNTANLTKSTFTFILLLLLRKKDQGLRFMDKET

#SKAFSFLRIEQENFSFKFKKSALSFTYNTANLTKSTFTFILLLLLRKKDQGLRFMDKET

LENIYRHFRYRFLKLNILPAFLGLLLLCSPNTLNYTQVDVIFSDRLCSCLLIFLAIASLT

KRSLLWLGAPLGIWTLFACVAGRSPTIFANDTLIGFAILAVVCISTRREALEVGPTLP

EGFSYNPSAGGRRAAVFLESLLGWLEARVLTASSLGITSSQSSNFLLLYSSIMTVYSLLV

VLSLAGSERRWHTRPKIVIATALALTGVIILTLPIILLQLRYDCWLCLCLTIEPALAVV

FAYDETRATLRYISQFLGDKRALTRASFFGSEYYKHTLSWEERTVLPLRKAYKQAFBGIS

FPINQLLAILVATVFVKVNSSMGLPTFPRNFLNICCWFIIVLFILAFAESLRHLRWMNLI

SAAAILFSPVLFHIPVESPMFLPIIVTGLILIILSIGKRRRTKRKL

CPn_0230 268277 267576
CT179 hypothetical protein
RFKKALIYMSSOPLVTTSSSLSRYVVLTGEEKVACYKKAFNHIWHGAPAIILAAALLMFC
IFGFVLGSILLGAPLEGASILYDVILFWLLPSILVFVLLVLPLNIYAYSHHKQVLALHER
ITOSNYKEIYDHCEKEKKTPNKKALSLYIESQVLVPEYSKRFSSMILGKTLKIIPKKDSP
ESLKHDELIQKALERAKENIYMNKNQREKRDEREAKKEAKNASKTNPLWEGLGT

CPn_0231 268996 268253
tauB-ABC Transport ATPase (Nitrate/Fe)
PQAFVSIQDAGFSHLQAHRLCYSCDNQVILKDASFQASPGTITIILGSSGVGKTTLFRLL
AGFLPLQBGELLWNSSPLMRKDVAYMQQKEALLPWRTALKNMTLSTELGINTSHNALSNE
RLEEIIHNFDLGQLLDRYPDELSGGQRQRIALAAQCLSLKPILLLDEPFSSLDVLLKEQL
YQDIVALAKKENKTVLLVTHDFHDVSCLGDVLYVIKNKTLTPVPLDPSMRPLNNGLCFIK
DLKKHLTY

CPn_0232 270134 269232
*similarity to 5'-Methylthioadenosine/S-Adenosylhomocysteine
Nucleosidase
KKFLMRRFLFLILSSLPLVAFSADNFTILEEKQSPLSRVSIIFALPGVTPVSFDGNCPIP
WFSHSKKTLEQQRIYYSCDSFGKYFVVSALWPNKVSSAVVACNMILKHRVDLILIIGSCY
SRSQDSRFGSVLVSKGYINYDADVRPFFERFEIPDIKKSVFATSEVHREAILRGGEEFIS
THKQEIEELLKTHGYLKSTTKTEHTLMBGLVATGESFANSRNYFLSLQKLYPPIHGFDSV
SGAVSQVCYEYSIPCLGVNILLPHPLESRSNEDWKHLQSEASKIYMDTLLKSVLKELCSS

CPn_0233 270439 270248
No robust homolog present in Genebank/EMBL as of 11/7/98
EKARTMFLGKVLLFLLRISRRSYVQEIGIFFHLETPDLKIVLCAFVSTFIVVEMDVSLKN
KGOS

CPn_0234 271246 270548
CT181 hypothetical protein
FIMLOSCKKALLSIVVSILAFHPIROMGVEAKSGFLGKVKGWFSKKEIQEEARILPVKDS
LSWKRYDYTSSSGFSVEFPGEPDHSCGIVEVPQSEITIRYDTYVTETHPDNTVYVVSVWE
YPEKVDISRPELNLQEGFSCMMQALPESGVLFMQARQIOGHKALEFWIVCEDVYFRGMLI
SVNHTLYQVFMVYKNKNPQALDKEYEAFSQSFKITKIREPRTIPSSVKKKVSL

CPn_0235 271395 272177
kds8-deoxyoctulonosic Acid Dynthetase
VFVRYLLMKPEESECLCIGVLPARWNOSRYPCKPLAKIHGKSLIQRTYENASQSSLLDKI
VVATDODHIIDHVTDFCGYAVMTSPTCSNGTERTGEVARKYFFKAEIIVNIQGDEPCLNS
EVVDALVQKLRGSFEAELVTFVALTTDREEILTEKKVKCVFDGEGRALYFSRSPIPFILK
KATPVYLHIGVYAFKREALFRYLOHGGTPLSDAEDLEOLRFLEHGGKIHVCIVDAKSPSV
DYPEDIAKVEQYITGLENAYF

CFI_0236 272138 273766
py_COTEP_Synthetable
py_COTEP_Synthetable
py_COTEP_Synthetable
performation of the company of the compa

VRGWEGK I AVAKFOREGG I PYFG I CLIGMQVLVVEYARN
VMEDQDPLVMTGGTMRLGAY PCLLKPGSKAHKAYNESS
IRGRI LVGTCPPQGLCE I I LEVDHPMM I GVQFHPEFVSKI.

PLF I AF I EAALVYSKDA

CPn_0237 273741 274214

yggF FAMILY GILRMOAMSKESSCKAYLGIDYGKKRIGLAYAAEPLLLTLPIGNIEAGKNLKLSAEALHK FIGURITY TYTEININGEROEST, ESTERIAEET SKILTFYETI DØDEFILIGVØAERM GREFF ISSURGERGEROETTSLAATELIDE FIGUREPKILT.

CPn_0238 274210 275838

zwf-Glucose-6-P_Dehyrogenase
PCNHQKLRDFNFRNFLLFVIFASAGTKKEIKMTNVVQETIGGLNSPRTCPPCILVIFGAT
GDLTARKLLPALYHLTKEGRLSDQFVCVGFARREKSNELFRQEMKQAVIQFSPSELDIKV
WEDFQQRLFYHRSEFDNNMGYTSLKDSLEDLDKTYGTRGNRLFYLSTPPQYFSRIIBNLN
KHKLFYKNQDQGKPWSRVIIEKPFGRDLDSAKQLQQCINENLNENSVYHIDHYLGKETVQ
NILTTRFANTIFESCWNSQYIDHVQISLSETIGIGSRGNFFEKSGMLRDMYQNMMQLLC
LLTMEPPTTFDADEIRKEKIKLQRISPFSEGSSIVRGQYGFGTVQGVSVLGYREEBNVD
KDSRVETYVALKTVINNPRWLGVFFYLRAGKRLAKKSTDISIIFKKSPYNLFAAEECSRC
PIENDLLIIRIQPDEGVALKFNCKVPGTNNIVRPVKMDFRYDSYFQTTTPEAYERLLCDC
LIGDRTLFTGGDEVMASWKLFTPVLEEWDQDSSPSFPNYPAGSSGPKEADALIERDGRSW

CPn_0239 275863 276672
devB-Glucose-6-P Dehyrogenase (DevB family)
KSISMTNIGIETMATLINFNDTNKLLLTKQPSLFIDLASKDWIASANQAIKQRGAFYVAL
SGGKTPLEIYKDIVINKDKLIDPSKIFLFWGDERLAPITSSESNYGQAMSILRDLNIPDE
QIFFMETENPOGAKKYQELIENKIPDASFDMIMLGJEDGHTLSLFSNTSALEEENDLVV
FNSVPHLETERMTLTFPCVHKGKHVVVYVQGENKKPILKSVFFSEGREEKLYPIERVGRD
RSPLFWIISPESYDIADFDNISSIYKMDIL

CPn_0240 277861 276698

No robust homolog present in Genebank/EMBL as of 11/7/98
LVYMVFSPSSESVVKANSVVRSNFCYFLENKFVSPSESTEVMFSEIMKGRVPDIESLFD
RPTDMMTGFKAAQNLGNLFNSFGILIMCFSQCKSCQTPEKETSAIVLGATLLFFVVALI
LGPTLGALVYCAYKVYTLGKMIYSLNKAKAKVLRHPAQNVFHRAAGVATIRSSEEAVKAC
KLYKSAMIGSLVVSLIASLALIALTAGIVLVLFFVAPGAAPVITAAMGCCAAGGGALLI
SLLGLWIAIVRKAKHQEACVGHLTNVVLHTAVSEALLHDPSHFQTNALARDLFLTDCLSH
YGHLFSNEEVAQLVQGGAPGGSRFSQHYGGSSDYQNRRGGNGNFGGSHFGGGGFAGSH
FGAGYPTAPTMPSAPPPFPPPAYDTIYG

CPE-0241 279372 278203
NO PROBUST HOMOLOGY DIESENT IN GENEBANK/EMBL AS OF 11/7/98
IFINKFMSAMISLSSSHEASIASNTQVRDVLVSLAMDEFVEHNTEILPIKVFLARGTLSS
TAJIDDLKDVVETEGEHHFQVYSNISLKMIYQRFFEKIFGIGCCFLLLVTDSHHTDPGGA
LITGIFAAVLFTVLAIVFGFTLGILCYSAXIYQLTKKISSLSRTHTEVINSVQKSDPFI
HRSGAVAAAASQSTIKACKVFRQSTLEFTVLGLITISLAALIVGLVFALFFLDPGAPA
WHTAMIGCCAAGGTGILLSVIGFLLASVYSVQKSQEGVHHMHTALLRGIVSNTIIQMPY
LPJEPGTKKVLTQSIRRYQQFFSDDEYRDIESEVPLNRQTTPPPSYETLFHEEGSDGSSN
VIPRESPPAYSTIDSSNSPFPSSSPPPYYR

THE CONTROL OF THE CO

Cri_0243 280609 280133
No Frobust homolog present in Genebank/EMBL as of 11/7/98
INPAYLVFLLKFVKGRIIMACSIGYHLCNANEPDRFVASKVALVADILLYPFMAVICAVV,
FAYLMYVKLLFLAIKFLVNTCIAACKSRPLPSCKEDFQCLFGPKDKPGPSDWLGCLVLIF
IIGFEIYSTIITVQSDTNRLRYFIISPAYQVGSTAIINW

CPHIG244 280906 281556
adk-adenylate Kinase
GARBATKGSVFIIMGPRGSGKGTQSQYLANRIGLPHISTGDLLRAIIREGTPNGLKAKAY
LDKGAFVPSDFVWEILKEKLQSQACSKGCIIDGFPRTLDQAHLLDSFLMDVHSNYTVIFL
EISEDEILKRYCSRFLCPSCSRIYNTSQGHTECPDCHVPLIRRSDDTPEIIKERLTKYQE
RTAPVIAYYDSLGKLCRVSSENKEDLVFEDILKCIYK

CPn_0245
281627
282499
ydhO-Polysaccharide Hydrolase-Invasin Repeat Family
TCOKEIMKHYLSFSPSADFFSKOGAIETOVLFGERVLVKGSTCYAYSQLFHYELLWKPYP
GHSFRSTLVPCTPEFHIHPNVSVVSVDAFLDFWGIPLPFGTLLHVNSONTYIFPRDILNH
MNTIWGSGTPQCDPRHLRRLNYNFFAELLIKDADLLLIFPYYWGGRSVHBSLEKFOVDCS
GFINILVOAGCYNVPRNAADOYADCHWISSFENLPSGGLIFLYPKEEKRISHVMLKQDSS
TLIHASCGGKKVEYFILEGDGKFLDSTYLFFRNNQRGRAFFGIPRKRKKFL

CPn_0246 282955 282551 rs9-S9 Ribosomil Protein VVAKSTIQESVATGRRKOAVSSVRLRPGSGKIDVNGKSFEDYFPLFIQRTTILSPLKKIT EDOSOYDLIIRVSGGGIQGOVIATRLGLARALLKENEENRODLKSGFLTRDPRKKERKK YGHKKARKSFOFSKR

CPn_0247 283430 282969
r113-L13 Ribosomal Protein
DDYTIMEKRKDTKTTIVKSSETTKSMYVVDAAGKTLGPLSSEVAKILRGKHKVTYTPHVA
MUDIVIVINAEKVRITOAKKOGKIYRYYTGYIGGMEEIPFINMMARKPNYIIEHAIKGMM
PROTELOKKQLKUSLRIVKGDSYETFESOKFILLDI

CPu_0248 D84453 283650
yetV/ybba ABC Transporter ATTuse
RDF0671224WR7ATEGREREDPACKERSKNIACHIFFORLAMGLLIEAKNLOKTIOOON
ON 15172707124WR7ATEGREREDPACKERSKNIACHIFFORLAMGLLIEAKNLOKTIOOON
ON 15172707124WR7ATEGREREDF7124CONSKTTELLHLLTTLDVPSCCSLREFDKDLKNODLA
NERGALIGEVEONFYLLEDDT7124KNVLMPALIARKYLISKGSPVYTRALELLDLVNLEDKV
RTRICOKLOSIGEPORVATAKALINETATLLADEPSCHILDEETGEOTHNLLLEOAGALCGTL
LVTIMKHLAGRCGRESTYLENGKLEFTING

CPn_0240 285850 28473 CP157 hypotherical profein LEVMKFEFSVALKYLIPGRO IVSLF*/GIISLVVWLSIVFISVIHGLEQR*IEDE SQLHSPITILPSDTYYSSTY HSSLS#ITTIKTLOEKIASPQVDFYDFESDYLLPET FPLKDCOLOGORDPVMTLESUGPYLOSØHGKVIEFEGGVDYLDIKTSLKLÖKROPRNL THFLTYPSKLSYEDKVLPYDETDYTSAELNPFNRSPSOWQOPFHLEELYRGASIILPST YKDSGYKVGDTGVFSTYSIENEKETQYTYHVIGFYNFOLSPLGGRTVFIDPDLARSIRSQ SEGLOMSNGFHLFFPNTKRIVFVKRQIENILTSLGVDDYWEISSLHODYPOPILDQLQS DQVLFLFVCILILIVACSNIVTMSHULVNNKKKIGILKAMTSSRSLKIIFACCGAFSG ACGVVIGTIFAIITLKNLOFIVKALNYLOGRETFNTAFFGONLPNSVHPQAIYFLGLGTL

CPn_0250 Zahdur 285332 r133-L33 Ribosomal Protein KDSSMASKNREIIKLKSSESSDMYWTVKNKRKTTGRLELKKYDRKLRRHVIFKEAR

CPn_0251 286336 287559

'conserved hypothetical protein
SPDSCLPWMSPFKKIVNRLICYISFOKESRTLPIIIREPRHTTKSLGSFNSVISKNKIHF
ISLGCSRNLVDSEVMLGIULKAGYESTNSIEDADYLILMTCAFLKSARDEAKDYLDHLID
VKKENAKIIVTGCHTSNHKDELKPWMSHIHYLLGSGDVENILSAIESRESGEKISAKSYI
EMGEVPROLSTPKHYAYYKVAEGCRKRCAFCIIPSIKGKLRSKPLDQILKEFRILVNKSV
KEIILIAQDLGDYGKDUSTDRSSQLESLLHELLKEPGDYWLRMLYLYPDEVSDGIIDLMG
SNPKLLEYVDIPLOHIADRILKGMRRTTSREGILGFLEKLRAKVPQVVIRSSVIVGFFGE
TQEEFGELADFIGEGFIDNLGIFLKSGEANTPAAELPDQIPEKVKESRLKILSQIQKRNV
DKHNOKLIGEKIEAYIDNYHPETNLLLTARFYGQAPEVDPCIIVNEAKLVSHFGERCFIE
ITGTAGYDLVGRVYKKSQNQALLKTSKA

CPn_0252 288112 287576
CT144 hypothetical protein (frame-shift with 0253?)
ATSTVCALWILDTYOSHDDAASCSFRACRFORYWLGGVNVPWNRFNGTSTQSTVINSAI
YIDSSQTWAMFGASASIFRLFRISIFMTKHGBUIDDATGGELLLVAYEANONPLFPDIR
IELAMSTCSGTSYYRARPMOWLCSTYYAVRPGYFVLENRSYSFRVQSFSWNIATLPFVN

CPn_0253 288474 287950
CT144 hypothetical protein (frame-shift with 0253?)
FCGGRLMSSSIPTTOKITISIPTFVRFNIESINLTDEOKKTALTIGONIATENTOVLGNF
VDADGGLICONDLSVGGNINITPOTFUTMFFSGRVALSNSPFSYODSLGNKRQDYANINT
EQPQQYVPYGYYKLTRVMMQRAALSGGHVGSGDIGWGESMYLGISSIKRQHKVQ

CPn_0254 289268 288459.
CT1/3 hypothetical protein
IPMKTLGVKDQNIFIDQATLSVERNVRIENNLETRDLKVLDTTTSPCEFIVKGNVSAEGS
QLANTTLSDGFNIYSKTDVSQTPVCNNISDPQSARDALTFSYYRKTGCQAANLYTYYPGN
GYYYAPNTTIETHVAAITSKSVSRNATPDFSRYADIEPVVKLKQVGIYQVTMQLTRMSGQ
UDGDNSATLILNFVSGNNKTLLCTSDTRGGYSSDRTSVAVTAIFSVTELVSSPPYDYPWI
ALESTIMMIMSLSTCVIWFPFPSNFVEVD

CPT-0255 290183 289329
CT14) hypothetical protein
TLLKYIMKNINNBECYFKLDSTVDGDLLAANLKTFDTQAQGISSTETFSVQGNATFKDQ
VSATGLTSGTTYNLNAQNFTSSQISIDFKNNRLSNCALPKEDCDPVPANYVRSPEYFFCS
KØLIGDFDFNSGESYLPLTGSEFTLYGSRNVNSIFRFIGWKQSTRELTVGGNTALQFLAA
GTYIVSFTVGKRWGNNNGGGAIYINNGLGQVQCESTIYSGGGYATGTGTGTSIYRASVD
VAPNPNDPNASDRYRAGIFYLSNGGSSAGIGNYSFSLLYYPDDRG

CPH_0256 291282 290398
CT144 hypothetical protein
FCGGRLMSNPTPKTKISIPTFVRFNIQSINLTEDQKKTTFTVGGKVTTENTVVRGDLTCT
DGGLTCQSDLTIQKDINIRPTSTNSMVFDGRLNLSNSPLSYKNSQGQDITDYEKMSSGKP
QEVVPFGYYKRTQIMAQRAAHSSGYVCGGSVPSGSYVPMNKFDQTSTOKTSGTEIYIDP
NDSTKLVFEVNNKVPKLFRISVIMAKHGSWLDNGTGADILLANNEYEGGGGRIWTDLAM
TTSRGSSYYETRPLQVVCVTYYAQNNGYFTFQNRAGGGLRVSFFSWNIVALPYVE

CPn_0257 292136 291267
CT143 hypothetical protein
GVVMKRRNLQKILPNASTPSTNVAENTGIKDONLFLDQATLNVDGNVDIENFLETRDLKV
ADTITSPCEFTVGGGLSAESSOFKATTLSKGLEITSEDQDGRVPKFTNVSDPQSPRDALT
YNYYRNTGCQALNLYTYYSSSQPTTVGKPIETVCONPNPETYRISASAKIYDAVTRFPYI
QFKAPGIYQVTIQIRRESGQHSGLDNPNLYLNLMIGNNKTLLCASDTRGYSGGHRTSIAV
TGTFTLTEIVATPPHDYPWLFLETTIGLDIKSMSTCVIWFPFQANFAEVD

CPn_0258 292534 292133
CT142 hypothetical protein (frame-shift with 0259?)
CFSFCRLGSKFERITLGGNTAIQLLAAGTYILTFIGRRWGWNNGWGGSIRLFEGKYTGD
GTMLCGSTVYSGGGYSTIGYLSTAVYRDHSDIDPDPNNPSDKYMNNFLFVRNGDHSAVIG
NYSFTLLYFAGDKV

CPn_0259 293031 292441
CT142 hypothetical protein (frame-shift with 0259?)
LYFVFKRKTYNYFIEMTTTMNQDNNECYFKLDSTVDGDLLASNIQTFDKQAKGISSTETF
SVQSNATFKEKVSATGLTSASTYKLNATGPAPSSITIGMKNNRLSNRALPKNPCDPVPAN
YVRSPQYFFCAKPIEGTFMFDGSSRYLPITGDGSNYTLYQSSKAGDVFRFVDWDQNSKKL
HLGGTQPYNFLLQEPIS

CPn_0260 294090 293548
secA-Protein Translocase Subunit
AYLDFSKRSCVEEDHVSKKINRIDLCPCGSNKKYKOCCLKKEEOTARYTTEGKFKFSAEV
LSASEQGEAGDNCTKLFQRLSOSLTSEOKAAVGKFHQITKNKEVMSKKALKKAQAKEEKL
VTEKLQOHNFEILNTGENLAPPMESTATLNQDTNFVCEDFIPTOEDFRISENSQKPPVEE
D

CPn_0261 294272 295033
yddo-PP-Loop Supertamity Atpase
YSPHINE FYPMCTLLLANPHMKACKR LESLVKALYTHTMLANHRK IVVALSOOKDSLTL
LLMLKA LOORGEFOLDLIAVN FORYSOOAEVINEYLTHE FODDLOL PERT LESPYAPETP
ECYTIGGARRELEFOAKE LOACALAFFIHRODLVOTALLILLIIKAEFACHLPVLOMVHF
GYT LLREL FEPPEFWLRKFAKEIKIFARVTORCEVVSLROYAEQELKLLEEVFPLARHNIA
LATGEHGOOKOOK L

CED_0262 295055, 295033 DBLE:GBLE: Like Acid Thoophataba LIENTRKEVKYYEVEMNKRUKT LEINDEX PEAKOMBOLVOALLEANTODIYTAAPQAEQO HEIMA EILNOVVCASPYAY POPVKEAWAVGGSPTDCVT LIFESVSPDLV INI TEKNAWYIXIT TGAAKQALVDG I PSHALSQONH I SF PELIKALVIY PELIKALVI I SPETOPOGSSWECHRLVPPTDEFFYEEPRYL NOVYVCKI SG GEELACMLENH I SVSP I FSONSPIGLITLEEFOKTQENFNASLLSSELTTKI F FESVSPDLVISGINCG PEILKALVIYLLSOPFP ONOYYVGKISGVRIGEHP

296174 297136

yqfU hypothetical protein STALSREKURVEPSLAKYAFROFRMSHGPRPTKFSFPLYFSKTLSWFILGGFLAACGVQ MODATE BOLE COLO DE LA CHELGHYAL DE LOVERNE PER ENTRE LOVIC GERE VICE DE LA CHELARA COLO COLO LA LA LA CALLO MODE POR POR EMPROPERTO DE LA CALVO COLO LE PROCEDE DE LA CALCA DE LA CALVO CALCA DE IIINKKKGYTVOOIILFVNFFIFALSGIVYKNWHTAFVSFLTYGIATKVMDMVILGLEDT KSVTIITSSPRKLGHILMETLGIGLTYIHAEGGYSGEPRNLLYVVVERLQLSQLKEIVHR

EDPSAFIAIENLHEVINGRRT

297730 297155

wbiD-Phenylacrylate Decarboxylase
MKRYVVGISGASGVILAVKLIKELVNAKHQVEVIISPSGRKTLYYELGCQSFDALFSEEN LEYIHTHSIQAIESSLASGSCPVEATIIIPCSMTTVAAISIGLADNLLRRVADVALKERR PLILVPRETPLHTIHLENLLKLSKSGATIFPPMPMWYFKPQSVEDLENALVGKILAYLNI PSDLTKOWSNPE

298632 . 297730 CPn 0265 CPH_UZO5

29852 297/30

ubia-Benzoate Octaphenyltransferase
KIIIVRLNYFLNLVNFKYSIFSILFLSASTVFALSINEISQNLSFKEGFKISVFGAIAFV
FARTTGIVVNQCIDRFIDKKNTRTSKRVLPANLVSLNFAWVLSLFCSFLFLELCKILRIF
SLGIASLTLMIVYPYMKRVTFFCHWGLGLVYTVAILMNFCAFAESGLSMRLCFLALLWG
SVGMVIAANDIIYAIEDTEFDREEGLRSVPAHYGEKKAVEIAKVNLWVSYLAYIFSGFVG

CPn_0266 299181 299876
No robust homolog present in Genebank/EMBL as of 11/7/98
IMALDEINNONNESQOIASSTSGTSKINODRRIFACTVYLLVVARLMILSGIVLLFTIGS
LGLSVPLSGILGTFAVTVGAVLFITGLITILVRKSLGIEQKNEDLMFLKIKTPTPPARPLM
SKFSVTCSTTSIVLGMALLIGAVVSVFFLTGYLQLGLCAGLVGLGTALFVAGLARMSPRS
LADQEGSGSADSQSNIVGIGEPKAAQEQKWYKMAVVRGEDGIPTAIRLTPEK

SLDKEFYFTAIIPLVVILKVVRMYSNYSKKDQEGESKFFLANIAIALSFLVSMTLFWSLS

CPn_0267 300122 300910 No robust homolog present in Genebank/EMBL as of 11/7/98 VSIMSLNKTNALLNOPEPAVCLNAWDPKYINOPRKTFACTVTLLVIATLMILTTGVIVLL AMGSPGLSVLVSTIIGTSVTTLGTALFIIGLVKLIKKSLAWIQYQKYFOEVVKOKYEPFS IPRNDNYHKLTSCLPSPLDIESPSPEASTPVSKLRIACSGVAIVLGVTLLIGAVVSVFFC
TCHEQLALCVGFACLGTALFVGGLAGLRTHSLIAQGIMYLYLTYYLSSALEERNETVKDQ RNEINTYLTEECROOKREKALLE

CPn_0268 300914 301318 CPL 10258

JUD14

JUD15

NOTFODUST homolog present in Genebank/EMBL as of 11/7/98

KQWALSLMSQCQSSSTSTWEMKKSFVPNWKNPTPPLSPIPSEDEFILAYEPFVLPKTDPE

NAQANPPGTSTPNVENGIDDLNPLLGQPNEQNNANNPGTSGSNPTSLPAPERLPETEENS

CP 0269 302468 301476

QEEEQGSONNEDLIG

Dipeptidase

VARRICVMTIDMHCDLLSHPHFCRKDPAVRCSPEQLLSGGVRQQVCAIFVPHSRGEPNCDK

QNSLFFSLPNQYPDIGLLSYEEEENGSSSQKKSLSLIRSIENASALGDDTAPLGTLLAKL

HLTKQGPLAYLGIVMKGDNRFGGGTEAPKRLSNDGKVLLDIMYELGVPIDLSHCSDKLA

EDILDYTADKLPNLAVIASHSNFRSVLDHRRNLVDAHAKEIVRRKGVIGLNLVRSYVGDS

LGDLEKHVLHAENLGILSSIVLGSDFFYANDEDENFFFNECSSAEAHPVLNQLIHRIFSKG KAESILSSRAEKFLKQVIVEQVNPKITDVKL

CPE-0270 303343 302468
ywlC-SuA5 Superfamily-related Protein
SIFGVIVPDKKAGITFSLPEVMSAINGGKIVALPTDTVYGFVLSLYASEAEERLYALKDR
EPSKAFALYVNSIEDIENISGYPLSPTAKKLAOLFFGAITLVVKHRNRFFRKETLAFFIV
DHSWREIVDHCOTLIGTSANLSEFPSALTAGEIFADFADHDLIFDGPSKGLESTVVA
SDPLYIYREGLISRSVIENIAGTEAKIFHRTSHAFSKHIKIYTVKNQEQLVSFLSGSLDF
KGVWCEHPKPKNFYTRLREALKKKTPSIVFIYDINTSDYPELFPFLSPYYIE

CPn_0271 303628 304362
Lysophospholipase esterase
KLMTDYSFFRRKIGNIEAIECFONPODPIIILCHGYGSLADNLTFFPSICSFSKLRPTWI
FPNGILPLENDFRGSRACFPLNVLLLQEISRLYANGVGNLQEKYDELFDVDLETPKEALE
ELILNLNRPYNEIIIGGFSQGAILATHLVLTSGNPYAGALIFAGARLFNQCWEEGLKQCA
QVPFLQSHGYEDEILPYHLGAHLNDLLTKLNGQFVSFHGGHEIPSVVFQKWQVTVPNWI
DPARG

CPn_0272 305272 304340
dnaX-DNA POl III Gamma and Tau
FNRQSDATYATWYMHLEEENQGWEALLRKVYHQEVPPAILLHGFTLPVLQDKAEQLASEI
LLSSSPOSEHKYSOKIHPDIYOFFPEGKGRLHSIDLPRGIKKQIYISPFEANYKIYIIHE
ADRMTLAAISAFLKVFEEPPKHAVIILTTAKVQRLPKTIISRSLSIFIERGEKILGSKET
FSYLFRYAQCEIPVTEVSQIIKESSETDKQVLRDKVQRFMEVLLELYRDRYTLNIGLKJKS ALNYPEHVKEILQLPLLPLDKVLLIVESACRSLANSSSAASVLEWVAIQLVSLQYKER VSVSPGODLSN

CPn_0273 305853 305227

CPT_0273 30593 305227

tdk-Thymidylate Kinase
GGIVFTVIEGGEGGGKSSLAKALGDQLVAQDRKVLLTREPGGCLIGERLRDLILEPPHLE
LSPCCELFLFLGSPAQHIQEVIIPALRDGYIVICERFHDSTIVYQGIAEGLGADFVADLC
SKVVQFTPFLPNFVLLLDIPADIGLQRKRRQKVFDKFEKKPLSYHNRIREGFLSLASADP RYLVLDAREGLAGLIDKVMLHTQLGLCT

308368 305852

GYTA DNA GYTGGE GUBURIT A ESTTIEMFRIKDETIVEKNUEEEMKESYLRYSMSVIISPALPDIRDJLKPYORRVLYAMKOL CLOSCARHERCAE POOTECOPYEINESS ERYSMYTTSPALPDERCHERYORRYCVAMKOL

CLOSCARHERCAE POOTECOPYEINESSYTYPTLYPMAQNAMARYELVOCONFOS LOOD

FINAME TEARLTHEAMYLMEDLOKDTVDIVPNYDETKHEPVVFPSKEPNLLCNOSSGIA

VOMATHI PPINID JELLEATLLLLANEQASVDETLOVME-OPDEFTTOG! CCCCEGIRSAYTT

GRICKIRVPAKLHVEREBOBRIERSI LITEMPYNNKOPLIEQIANLYNEKTLAGISDVRDE

CLKLOTEVAVLEIKKGESSEI LITEMPYKFTDVQVTEGAIMLALDKKLPRTMG IHRMISAWI

RHKKEVIPPKTRYELAKAETRAHVLEGYLKALSCLDALVKTIREGGNKEHAKERIIESFO TTEPOALAILELRUYQUTY COMEYEELLNKIAYYKOVLSDEGUYKDI IRNEQDL LKHHKVARRTTIEFDADDIN ITNEGVI ITISGDDYVKRMPVKYFKEORROGHGYT GFDMKKGAGFLKAVYGAFTKDI LLITTNEGOCYMLKVWOLPEGERBAXGKPI INFLEGIR GEELAALINI IKNFONAGFLELATKRGVVKKVSLDAFSNRKKGIRALEI DEGDEL IAAC HIVSDEEKVALFTHLGMAVFFHEKVRFMGKTARGVRGVSLKNEEDBVVSCGIVTENGSV LIVCDQGFCKRSLVEDFRETNRGGVGVRSI LINERNGNVLGA I DYTDHDSILLMSSQQA IRINMODVRVMGRSTOGVRLVALKEGDALVSMEKLSSNENDDDVLSGSEEECSDTVSLR

310 744

PTD_TOTAL SYLENGER B FINDPRENT BRANCE FROM TO THE THAT CONTROL OF MGEEVPPRREFIESHALSIRINNLDI

31/0793

CTI91 hypothetical protein /
DMFLKRKKRGGSQVONKRTASPIKHAKYYLHNYLQELQKIMAARPHDAIDAWNQVFRDKY
KGMSQAIGFRDHILLVKVYNSSLYALEKQTPQNDLIMSLYQVASHVQIREIQFLLG

CPn_0277 312003 311404
No robust homolog present in Genebank/EMBL as of 11/7/98
NISIFYPKYFIEGKEVLIKNLPPLFYGVILMIINVRAPAFGITSVQOFSTNFQAAIPIL
NIVIGCSRISSTYAEDIEEVAQEKLEKSTHSKSSTSVNLWAHRVRGVVEILGGGIVILAL
EITALVLQVIIKLIKCLIDVLC/CLFGLGVCVVAIIGAIAFCVVVVVKYLGFCSQGEELE
PIEVKTLISPDKPYPTVVVV

31/2884 312060 CPT_U278

39284

31200

**Conserved outer membrane lipoprotein

RDSMKKKLSLLVGLIFVLSSCHKEDAQNKIRIVASPTPHAELLESLQEEAKDLGIKLKIL

FVDDYRIPNRLLLDKQVØANYFQHQAFLDDECERYDCKGELVVIAKVALEPQAIYSKKHS

SLERLKSQKKLTIAIPVBRTNAQRALHLLEECGLIVCKGPANLMTAKDVGKENRSINI

LEVSAPLLVGSLPDVJAAVIPGNFAIAANLSPKKDSLCLEDLSVSKYTNLVVIRSEDVGS

PKMIKLQKLFQSPSVHFFDTKYHGNILTMTQDMG

CPn 0279 313546 312875

CPn_0279 313546 312875

* Possible ABC Transporter Permease Protein
KKIMOSDLIQILKETVNILMVSTAFFFSCAIGGMLGLGLFCTSPKSLNPKKSLYATIS
MILSFLTAIPFAILIVILFPITRWIVGTSLGPTASIVPLTIGAIPFVVTIVVDAFRNSAL
NYLESAVALG PERNILEGILLPESYPOLIFSLKSLVVHLISCSTLAGFVGGGGLGQLLL
QYGYYR REASYTTSVIDITLVLIESVRILGDFWGRRVLKYRGIL

CPn_0280 314593 313550
dppF-Dildeptide/Transporter ATPase
IKGEAWLVSEDHSFIISVDDVSKKLGDHILLSKVSFSVYPGEVFGIVGHSGSGKTTLLRC
LDFLDMFTSGSISVAGFDNSLPTOKFSRRNFSKKVAYISQNYGLFSSKTVFENIAYPLRI
HHSEMIKSEVEGOVOTTLNFLNILYHRHDAYPGNLSGGGKGKVAIARAIVCQPEVVLCDEI
TSALDPKSTEM; IERLLOLNGERGITLVLUSHEIDVVKKICSHVLVMHGGAVELGTTEE
LFLYSENSITNELFHEDINIAALSSCYFAEDREEVLRLNFSKELAIGIISKVIQTGLVS
INLLSGNINLFRKSPMGFLIIVLEGEVEGRKKAKELLIELGGVIKEFY INILSGNINLFRKSPMGFLIIVLEGEVEQRKKAKELLIELGVVIKEFY

Cpn_0281 n_0281 315033 316103 ChnA-Predicted 1.6-Fructose Biphosphate Aldolase (dehydrin

JEMBILY) I SIRRHTIMENTHDILGNDDENLLSYQCKHITKDKLTLPSHDFVDKVFGLSDRNNRVLRS LOTMFSHGRLANSGYLSILPVDQGIEHSAGASFAINPIYFDPENIVKLAIESGCSAVAST YGTLSLLSRKYAHKIFFMLKLHNELLSYPTKYHQIFFTQVEAAYSMGAVAVGATVYFGS ETSNEEIVAVSNAFAKARSLGLATVLWCYLRNPAFVANGVDYHTAADLTGQADHLGATLG ADIVKQKLPTCQGGFKAINFGKTDERVYSELSSNHPIDLCRYQVLNSYCGKVGLINSGGP SGKNDFTEAARTAVINKRAGGMGLILGRKAFQRPLSEGIQLLNLVQDIYLDPNITIA

CPn_0282 316084 317529

xasA/gadC-Amino Acid Transporter
ILILOSLNFSKKVPHHSHSKPTKPLGTFTVGMLSLAVVISLRNLPLTAKHGLSTLFFYGL
AVICFMIPYALISAELASFKPOGIYIWARDALGKWMGFFAIWMOWFHNMTWYPAVLAFIA
STIVYKINPELAHNKYYIATVILAGFWILTFFNFLGITSSALFSSICVIIGTLIFGOILV
SLALFWIFSGNPIAISLSWGNLLPNFSNVSSLVLLAGMLLALCGLEANANLASDMVNPRK
NYPKAVFIGAIATLTILVLGSLSIAIVIPKEEISLVSGL/KTFTLFFDKYNLSWMTGIVV
WMTIAGSLGELNAWMFAGTKGLFISTQNDCLPRLFKKVISKNVPTNLMLFGGIVVTIFTL
LFICLDSADLVYWILTALSVGMYLAMYICLFLAGPILRIKEPRAQRLYSVPGKFLGICTM
SILGILSCAFALWVSFLPPRELAQISEGSKIGYTTFLLLAFSLNCLIPFGIYFTHKRLSK
KS

CPn_0283 318581 317532
No robust homolog present in Genebank/EMBL as of 11/7/98
GRRLGYFQDLIKNAVAKIISFRKSPPNEVKLLIKFAKKDLIKNSSIAFLYEVLLEILEAPG
EFILEVLFSLDPMULKSMLDFKKHSTLGIETSSFTAETIESSGLGLISINLLISGLCLRS
SHORGOAVKIIQOFCTOFFSSEEVONFVEORNILTPFLHHLFEGDEVALLNQLGURLDLIV
PNALVFEPDPSCWGINSETCAKDAEDQGEDFNKTKEACKEBLKKLVLPALGITSIFQLL
RARRFKQGAEILMAITRKKMKONFFIFLEALLEGEETIGVGKYLKLLMTHLWDKLLHA
LYLGYFOYGLIGOGELETFCRRANINPEAFOAAIOOORLLGFLFFKMLLD

CPn_0284 UP054 318551
No robust homolog progent in Genebank/FMBL as of 11/7/98
FLIMBLEPAPQVPVIDEPVNNTSSYGESTROOMERTYLITATIATATATHEVLYFGGTIS
WITHYLIMBLETSWIGGTAVAMIEYONGSTERTRYGSTTORSVERSDEPLNILLIGREEDS
VSAIDELLKNFPABPERRPKMLYCHFLDEOGRPNESPEEDSHTSKIL

CPH_928'S COUNT 119051
NO robust homotos present in Genepank/FMFH, as of 1177796
KECFNLFFFTANKETFASHELTYEKHOCFSLSP/TIL/SLAATSVELLESVVFALWGERVI.

AAPLGLLVWGCAASVCSMMAIVSLMCLYKGGKPLIEPS
PVEQGSLPKERKTVSFKAKIPSIVEDDFKPYVIGSTF
ILIVDFPRALEESSKSSGSLLAGVISEIKNLFLPRFLS*
**SITACLPRLGSIVEEYA
**SSOLLILLITKPEPLIMVTQOLIAHLINSLKTEKRKLTPHMOKUVLSINFWFYGMSLEKE
**IEKIVAYDPNLLTDELKAHLEAGNIVOFLLSFQSSEMGREFRALFPSDAQELPSAKDGSN
**YVPAINSSEYMTDFKDLSVKKKLSERLAFCEKIPSPSSWNFTSSVASHYKDFSLLFTFF
**SNQSVILONPFLLIELLHENPKCOTFLKGLLEKAMPMSNWAALFRPMLMGMLCSGIARK
KELKIIAEHLTVPFKEITOAIASGKILDLLLQHLFDF

mgtE-Mg++ Transporter (CBS Domain) 320650 mgL-mg-+ Transporter (CBS DOMAIN)
SCRESKGKIMMGEDNRHEEKLDTAFSSGRLMDSRTSHLDDELSFKLEKAFTCLSTDIHSH
DLSKIVIEYNPIDLAYAVSCLPSESRAILYKNLSCITAKVAFIINTDSASRWAIFRRLSD
SEVCALIEOMPPDEAVWVLDDIPDRRYRRILELIDSKKALKIRDLGWIGRATAGRLMTNE
FFAFLMETTVKDVSACIRSNEGIDLTRLWFVLDFKGELGGVVTDRSLIINPPEMSLKQIM
NQIEHKVLPDATREEVVDLVERYKIAALPVVDEENFLIGAITYEDVVEAIEDIADETIAR MAGTTEDVGYQTCHVVQRFLLRAPMLLVTLFAGLISASVMAYFQKISPALLALIIFFIPL INGMSGNVGVQCSTILVRSMATGTLSFGRRRETIFKEMSIGLLTGVVLGILCGLVVYLMG FLGLNIFSGGGIQLGVTVATGVLGASLTATTLGVLSPFFFAKLGVDPALASGPIVTALND IMSMIIFFLIAGGINFLFFN

CPn_0287
324230
322089
No robust homolog present in Genebank/EMBL as of 11/7/98
RRCMIRSPLPFISSKRALNMLGLQDEFSCPEDVVDFLFSEIELLAQQDEPSECYLALSRS
LLMMTHNHPKVVKRVIFYGVSYGLKHKSMSIFIDVLTYIDFLFEKLGISASDRLSLCSAR
TCINFELYSQTGEMKFLSEVVDNFRLIEGLLKMHPDLKNRLGMEHFRIGAKQEEVSLVAS
ASVYQAVGRSFIELYHKHLELSDLACGMKCLALALDLSPNNARIHADYAKGLUVLGTRQG
KSLLIERGMEHFSKAIFLSFSRDGDTLAYQNYRYSYALASVKLFDLTYKKEHFDQAMNIL
YQTVQAFPNLSGLMWVMGELITRSGLMSNMMYIEVGLEKLASLQKKTNDPIALSGLAT
GIAILGLYLEEPNLFKDSRHRLISAMRTFPGNSALVHALGVVQLCSALYFNEDSHFASAI
SCFQSCLEWDLDATGMWQKLFDAYFSWGIKKKSARLLRKAVDVASRLCSLRPEAFLFWSD
RGLALKCLAEATIDFAYKEIFLSESLHYQRAWDLSGRLEILELMGQSHYLLAELQQSLF
HYDEAYTLLTKVDLTLSSSRVKLILAAVLLGKGRLLQDTDPAEEAREILEPLVEVYLEDE
NFLLLLGKVYLFLFWKNRNVCLGKLARTYLEKATSLGCPEAYYTLGKFYAVIKDVNKAWG
MVIRSAQYGVRITEAKWLNDPYLANLREIHAFREVVENQKGRLWLGNKTEMKRN

325785 324571 CT288 hypothetical protein
ISITIREFLFFGFECRAKFYNVIMSCFNLTSTNESLRPISPKASFPKQGWQSYFRSALRK ISTITIREFLEFGFECKARFYNVIMSEFNLTSTNESLRFLISFAASFRAAWSIFKSALKA HESETLSSVCKVNKKVANDALFFNLTVIALAVVGVLILFS HILASIGGTLVITSFBELVTAA IËLFTILLTGGMYILHRLGKKVDVISGVCIPPFSRRCWVPISSSHTLEKFDEKHVSACSY LDJSTLSADGSGIAAVYQCPPLLFRAFPCFGIPCAMPFVALLRMTYNLIRFLWYPFYIIF RMFYEFFFCKHLPEDDRFTYKDVAREWGRSLAAFIKAPFYASACMIGAFYSLLDPLAGRV LMGSVERDWNDNVILARSVSLANEAHSLFRFEGGGGRKGLGQHAFYLMLCCQPQSVFLFD KGETVSGAHPSIQLPERGLDTSGRYPHISVIPDSGNDSAKNFIV

325797 326996 CTEE9 hypothetical protein NEWELMKKORSHYKKNNLLLLLSILVGLGLGSVQSPWIVYSAECIANTFLKFLRLLSIPL NENGLARAGORSH KRANDLUSSILVGUSGSVOSEWIYSAECIANTERFEREISSPE VEČALGSTITSIONFNTMYTLGKRILYYTLLTTVIAASIGLILFFLERPOHTODALATT TKENPLGYLDVLSDTLPENIFRSPELOGNVISAACLAVLLGTASLFLQEKEKHFVNQFFNS FESIFINLARGGLKLLPIAMLGFSVILFKELKDGSNLTHFAEYLLCVIGANLAGGFIVLP IGERINKVSSPLKVAKAMSPALVTAFFSKSSAATLPLTMELAEDDLKINNNLSRFSPFLCS VENDACCAAFILTTVLFVATSNGMIISPLMSLGWIFIATLAAIGNAGVPMGCYFLTLSLL TSMNVPLSILGLILPFYTVIDMIETSLNVWSDCCVVSLAN

327027 328523 Na-dependent Transporter RSALTMNKKHASFSSRLGFIFSMIGIAVGAGNIWRFPRVAAQNGGGAFLILWLCFLFLWS IPLIIIELSIGKLTKKAPIGALIKTAGKKFAWAGGFITLVTTCILAYYSTIVGWGLSYF YAYSGKIHLGNDFAKLWTSHYQSSIPLWAHLTSLGLAYLVIRKGIVHGIEKCNKILIPA IMPOGRIFILEND FALMTSHYQSSIFUMALITSLGLAYLVIRKGIVHGIEKCNRILIPAF FLCTIALLLRAVTLPGAVQSIKQLFSCDRSCFSINYKWIEALTONAWDTGAGAGGILLVA GERŠKKTGVVSNGALTAICNNLVSLIMGIIIFSTCASLDILGTTQLQDGAGASSIGIFFI YLEGLFTRLPGGIYLTTLFSSIFFLAFSMAALSSMISMLFLLSQTLAEFGIRPYISFILA TTÉMÉFVLGIPSALSLTFFSNQDTWGVALIVNGLIFIYAALVYGFPKLKKEVINAXPGDL RLWKAFDYIIKYLLPIEGILLLGWYFYEGLFPENGOWMPISLYSLGSLVLOWSLGLIIL WKENKQLYLRFSRYNHEIL

CPn_0291 328658 329194
incB-Inclusion Membrane Protein B
EKHMSAPIPTPQELSDQITCLAVQYQOVSELARENKGDIEGLKTLTAALTADAGIQPSAD
EIYSLQTAAALILSASEKROSGPSGSTEGSVTVQSFCKFKKVLAVVLTIIJLIATAVLIA
CIIAACGGFPLLLSALNLYTIGACVSLPIIASTSVALICLCTFVANSLIYPVITVRTTR

329201 CPD_0292

J29201

J2985b
incC-Inclusion Membrane Protein C
VKNTKNSDFMTSFIPPQSSCDASFLAEQPQQLPSTSESQLVTQLLTMKHTQALSETVLQ
QQRRLPTASIILQVGGAPTQGAGAPFQPGPADDHHHPIPPPVVPAQIETEITTIRSELQ
LMRSTLQQSTKGARTGVLVVTAILMTISLLAIIIIILAVLGFTGVJPQVALLMQGETNLI WAMVSGSIICFIALIGTLGLILTNKNTPLPAS

CPn_0293 329940 332723
CT234 hypothetical protein
WSMORVLRLENLHHGEEKRAFLFFLLGLVWGIGCYGTLSLAEGLFIEKLGSAELPKIY
LGSSLICVLSSLILVNLFKKHISATALFLIPVSLSILCNFYLILSSIFAIDPRSPLFF
YRIVIWSLTILSYTSFWGFVDQFFNLQDGKRHFCIFNAIJFLGDAIGSGIIASLVHTIGI
OGILLIFTAALUTTFPIVFYVSKSLKSLSDDHDLFIDTGHPPPLSKALKLCFYDKYTFYL
LCFYFLMQLLAIATEFNYVKIFEIOFASKEFEELVAHIJKKSLWISIGGNOFALFAYSTL
VKRLGVNNIILFAPLAFLSLFLFWFKTTLSIAVLAMVREGVTYALDDNNLQLLIYGVP
NKIRWJRI IVVESFIEPIGMLVWGLICFLSSQOYVFGLIISLIATILVCLVRGYYAKAIL
KNLGAOALOLTRSMODWIKOMTVKOKROVELFLLAHLKHPSERHOTFAFOHLLNLASRGV
LPSLLAHMNKLSLFNNLKTIEMVKSSLWAKDFLTLJLLKRWTSIFPHPAIAGAIHLYFAR
LDLLHITHIAEDLYDTVGDRLLAA LTVYRGGAVYDFVBLLADKRLKELBNSOPPEDIVAM
LTILKLEKNPONFPILLDFLNTKNEDILITTSQLKNTGVRANHKPYCFELLKRLRCGSHN
DEAGOYLLKTIGIALDIDFVKDLWTTTSQLKNTJKRYARAMIGELDKEVAPAFLOVLTDE
THENEGRILAAKALCKIONWLLKKHAYKIVKGJECKALFYSVIGHYIQKKYPTYNLSLLA
NIILENYYARAVMFMISILGIILGSMEHSGVLIFFALTSKNOKIKAQALESLEKHUZDSHLFSL
CEPFYNORISMCYSEKYYRKGGVIPTLEKELLMMMENSPOSLNKUTAQQLKEELSYCDPDF
OGVNTIYNOGHEDFRTEESETLISFLSI

CPn_0294 333077 11/502 CAMP-Dependent Prot Communication (Subunit Information Prot Communication Communicatio **ECPSVALSFLELYAKQIKFREP**

133627 CPT_0295
33386 /333627
acpP-Acyl Carrier Protein
AMSLEDDVTAIIVEQLOVDPKEVNENSSFIEDLNADGLOLTELIMTLEEKFAFEISEEDA
HELPTMIDVETTIEKPROAEC

334022 CPH_020b 33470£ 334022 CT296 hypothetical profein KIPIRGMICHDITLVGKKVIV/GGSRGIGLGIVKLFLENGADVEIWGLNEERGQAVIESL TGLGGEVSFARVDVSHNGGVØCVQKFLDKHNKIDILVNNAGITRDNLLHRMSEDDVS KSTNLTSLYYTCSSVIRHMIKARSSIINVASIVAKIGSAGOTNYAAKAGIIAFTKSLA KEVAARNIRVNCLAPGFIEFDMTSVLNDNLKAEWLKSIPLGRAGTPEDVARVALFLASQL SSYMTAQTLVVDCGLTY

CPn_0297 335724 334774
fabD-Malonyl Acyl Carrier Transcyclase
SRSNRDDNFMKRYAÐLFPGQGSQVVGNGQDLYMEYPEVRELFDFANERLGFSLTSIMFE
GPEDLLMETVHSQLAÝLYLHSMAVVKVLSQRSSIQPSLVSGLSIGESTALVASDRISVLDG
LELVRKRGQLMNEAKNQSPGAMAALLGLPSEVIEENITSLGGGIWIANYNAPKQLVVAGI
AEKVDQAIELFRDJGCKKAVRLKVSGAFHTPLMQVAQDGLAPDIYALCMKDSSLPLVSHV
VGKSLVNTEENRÐLARQMTSPTLWYQSCYHIESEVDEFLELGPGKVLAGLNRSIGISKP
ITSLGTFAQIEKFLSEV

CPn_0298 336742 335717 CPn_0298 336742 335717
fabH-Oxoacy1 Carrier Protein Synthase III
YTSFFLYMWFSVNNNKKAAIWATGSYLPEKVLSNADLEKMVDTSDEWIVTRTGIKERRIA
GPQEYTSLWGAIAAEKAIANAGLSKDQIDCIIFSTAAPDYIFPSSGALAQAHLGIEDVPT
FDCQAACTGYLYGLSVAKAYVESCTYNNVLLIAADKLSSFVDYTDRNTCVLFGDGAACV
IGESRPGSLEINNLSLGADGKLGELLSLPAGGSRCPASKETLQSGKHFIAMEGKEVFKHA
VRRMETAAKHSIALAGIQEEDIDWFVPHQANERIIDALAKRFEIDESRVFKSVHKYGNTA
ASSVGTALDELVHTESIKLDDYLLLVAFGGGLSWGAVVLKQV

CPn_0299 336726 337415
reck-Recombination Protein
RKKLVYYSESLYSNINIGPRPECKNKIHITHTRYPDYLSKLIFFLRKLPGIGFKTAEKLA FELISWDSEOLKILONAFHNVASERSHCPLCFTLKESKEADCHFCREERDNOSLCIVASP KDVFFLERSKVFKGRYHVLGSLLSPITGKHIENERLSILKSRIETLCPKEIILAIDATLE GDATALFLKQELQHFSVNISRLALGLPIGLSFDYVDSGTLARAFSGRHSY

337768 340152 yaeT-Omp85 Analog GRLLGMLIMRNKVILQISILALIQTPLTLFSTEKVKEGHVVVDSITIITEGENASNKHPL GRLLGMLÍMRNKVILQÍSILALIQTPLTLFSTEKVKEGHVVVDSITIITEGENASNKHPL PKLKTRSGALFSQLDFDEDLRILAKEYSVEFKVEFSEKTNIALHIAKPSIRNIHISG QVVPEHKILKTLQIYRNDLFEREFFLKGLDDLRTYYLKRGYFASSVDYSLEHNQEKGHI DVLIKINEGPGKIKQLTFSGISRSEKSDIQEFIQTKQHSTTTSWFTGAGLYHPDIVEQD SLAITNYLHNNGYADAIVNSHYDLDDKGNILLYMDIDRGSRYTLGHVHIQGFEVLPKRLI EKQSQVGPNDLYCPDKINDGAHKIKQTYAKYGYINTNVDVLFIPHATRFIYDVTYEVSEG SPYKVGLIKITGNTHTKSDVILHETSLFFGDFFNRLKLEDTEQRLRNTGYFGSVSVYTVR SQLDPMGNADQYRDIFVEVKETTGNILGIFLGFSSLDNLFGGIELSESNFDLFGARNIFS KGFRCLRGGEHLFLKANFGDKVTDYTLKWTKPHFLNTPWILGIELDKSINRALSKDYAV QTYGGNYSTTYILWEHLKYGLFYRGSQTSLHEKRKFLLGPNIDSNKGFVSAAGVNLNYDS VDSPRTPTTGIRGGYTFEVSGLGGTYHFTKLSLNSSIYRKITRKGILKIKGRAFTKPYS NTTAEGVPVSERFFLGGETTVRGYKSFIIGPKYSATEPQGGLSSLLISEEFQYPLIRQPN ISAFVFLDSGFVGLQEYXISLKDLRSSAGFGLRFDVMNVPVMLGFGWPFRPTETLNGEK IDDVSQRFFFALGGHF IDVSORFFFALGGMF

CPn 0301 340163 340762 CMT_USUI 340163 340762 (OMDH-Like Outer Membrane Protein)
IKOLSKEIFVVFRKGFWYFFSIFKLVQVIMKKLLFSTFLLVLGSTSAAHANLGYVNLKRC
LEESDLGKKETEELEAMKQOFVKNABKI EEELTSIYNKLQDEDYMESLSDSASEELRKKF
EDLSGEYNAYOSOYYOSINOSNVKRIOKLIQEVKIAAESVKSKEKLEAILMEEAVLAIAP **GTDKTTEIIAILNESFKKON**

340766 341866 CPn_0302 340766 341866

IpxD-UDP Glucosamine N-Acyltransferase

SKFKEFSMSEAPVYTLKQLAELLQVEVQGNIETPISGVEDISQAQPHHIAFLDNEKYSSF

LKNTKAGAIILSRSQAMQHAHLKKNFLITNESPSLTFQKCIELFIEPVTSGFPGIHPTAV

IHPTARIEKNVTIEPYVVISOHAHIGSDTYIGAGSVIGAHSVLGANCLIHPKVVIRERVL

MGNRVVVQPGAVLGSCGFGYTNAFGHKPLKHLGYVIVGDDVEIGAMTIDRGRFRVTV

IHEGTKIDNQVQVAHHVEIGKHSIIVAQAGIAGSTKIGEHVIIGGQTGITGHISIADHVI

MIAQTGVTKSITSPGIYGGAPARPYQETHRLIAKIRNLPKTEERLSKLEKQVRDLSTPSL

AFIDSFI

CPn_0303 342982 341921
CT303 hypothetical protein
REQKGLHHMDVSRKINRHTQFYVDSIDGVIKNFDHKPSEDKSRDHEELEEKLLTITKRIV
ASAQEFONRKTDSKNYYLKKTQWLPFKNEELEQTKELFAMLTSMDKKIAQLFFYSPGCSS
DWVEFTEVICHLNDSIGLGGVLLCGGLEEQQCEHVVTVNKKLDLPLLLGTTVVNSLRYYL
TYRNISLLNCQSMSELGKELGDVLKQHGVAFTLIFKEIVDIDLLNYVKLIQGLKRSGNIQ
ARIYDNDVPTLPSVSSSPIALRYSLANTIRGLALHVDFSSLKFISPSILSNTEHTAKALN
SGGECFIFSNLDEFNLGMKIVMQLLRTGKISPEILNKNIMKILMIKRRVRSLYI

343091 344158 pdth/odpA-Pyruvace Dehydrogenuse Alpha DQKPLPKRLPYKKVMDSSAPYNIASOYTEKSTVERILDLYGPASCIKFLKOMVLIREFEA DONEUPRICE TROMBOS AGENTACION TERROT VER LIDDEVO PARO LIFEROM VEI REFEA
RGERAY LEGELVGO PHISTA CORANTA A LANTI LEDDEVO PAS Y PICHALA LILLATI FLOETA
ELLAKET CALARCOS MINICOPPETRIEG LVCRO LIPLAGA A PTI KYOFOKNRYS LCPIO
DONAC A VEHIET LINEVELLO LEGEL LEGEN MOMOTESI A REVAKO LI A ECCOSTO DI RAV
TYRA CELE PRESILA PRESANTA MATERIEN PALORELI CODDILIVACIO ENEMNO CIPEK
DPI VLAPUMELI RIEVUTE E EPONI ROPCKTA VLEA PONAKLOODE SYTTLE ECVA.

144142 CPO_0.305 STATEMENT OF Dehydrogenatic Botto Botto Between Betw

345136 CPT_0306
pdnC-0ihydrolipoamide Acetyltransferase
GKFVIGLLKMPKLSPTMEVGTTVKWHKKSNDQVSFGDVIVEISTDKAILEHTANEDGWIR
EILRHEGEKIVIGTPIAVISTEANEPFNLEELLPKTEPSNLEASPKGSSEEVSPATTPQA
ASATFTAVTFKPEPPLSSPLVFKHVGTTNNLSPLARQLAKEKNIDVSSIQGSGPGGRIVK KERRYALTETTIA JE TYTETIETZETTÖTHEENIJE FEFT FAARLIAARI (LITETTYRIQQVY ASITARIILIETELJALTIKI, TITETTYRAISIALKEETTI NOOFTOTTAK I VEFTTI LI LI IA VAIPOGI ITPI IRCADRKNIGMISAEIKSLALKARNQSIQOTEYKGGSFCVSNIGMTGIT EFTAIVNPPQAAILAVGSVTEQALVLDGEITIGSTCNLTLSVDHRVIDGYPAAMFMKRLQKILEAPAVLLLN

348998 346515

GPI_JUTY

glgP-Glycogen Phosphorylase

NGCIVEDFSSFDKNKVSVDSMKRAILDRLYLSVVQSPESASPRDIFTAVAKTVMEWLAKG
WLKTQNGYYKNDVKRVYYLSMEFILGRSLKSNLLNLGILDLVRKALKTLNYDFDHLVEME
SDAGLGNGGLGRLAACYLDSMATLAVPAYGYGIRYDYGIFDQRIVNGYQEEAPDEWLRYG
NPWEICRGEYLYPVRFYGRVIHYTDSRCKQVADLVDTOEVLAMAYDIPIRGYGNDTVNSL
RLWQAQSPRGFEFSYFNHGNYIQAIEDIALIENISRVLYPNDSITEGQELRLKQEYFLVS ATIQDI IRRYTKTHICLDNLADKVUVQLNDTHPALGIAEMHILUDREELPMQTIFDVS
ATIQDI IRRYTKTHICLDNLADKVUVQLNDTHPALGIAEMHILUDREELPMQTAKMENTT
VIFNYTNHTILPEALERWPLDLFSKLLPRHLEIIYEINSRWLEKVGSRYPKNDDKRRSLS
IVEEGYQKRINMANLAVVGSAKVNGVSSFHSQLIKDTLFKEFYEFFPEKFINVTNGVTPR
RWIALCNPRLSKLLIVETIGDRYIIDLSHLSILTSFAEDSGFRDHWKGVKLKNKKODLTSRI
YNEVGEIVDPNSLFDCHIKRIHEYKRQLMNILRVIYVVNDLKENPNQDVVPTTVIFSGKA APGYVMAKLIIKLINSVADVVNODSRVNDKLKVLFLPNYRVSMAEHIIPGTDLSEQISTA GMEASGTGNMKFALNGALTIGTMDGANIEMAEHIGKENMFIFGLLEEQIVOLRREYCPOT ICDKNPKIRQVLDLLEQGFFNSNDKDLFKPIVHRLLHEGDPFFVLADLESYIAAHENVNK LFKEPDSWTKISIYNTAGMGFFSSDRAIQDYARDIWHVPTKSCSGEGN

CPn_0308 349213 349596
No robust homolog present in Genebank/EMBL as of 11/7/98
FFTOENMATVACTPOTTOPOPSVSHKATHRYCSWFFKPILVSLGLLLASLTTLGLVIA
SGVTLSLGIGIVLAIQIVLAGIALVLAFNHIRQFKQARTAELNSMKMISAPAAATVOKQK

CPn 0309 350977 349595

CPn_0309 350977 349595
CT309 hypothetical protein
FMRAWEFILLQEKEIGTINTUKWLRSLKVLCFDACNLYLEAQDSFQITWFEEHIRHKVK
SGLWNNNKPIRVHVTSUDKAAPFYKEKQMQOEKTAYFTMHYGSVWPEMTFSNFLVTPEN
DEBFRVLQEFTKSPDENGGVTFNPIYLFGPEGSGKTHLMQSAISVLRESGGKILYVSSDL
FTEHLVSAIRSGEMGKFRSFYRNIDALFIEDIEVFSGKSATQEEFFHTRNSLHSECKLIV
VSSSYAPVDLVAVEDRLISRFEWGVAIFIHPLVQBGLRSFLMRQVERLSIRIQBTALDFL
IYAESSNVKTLLHALNLAKRWYMKLSHQLLYEDDVKTLLEVDLEAGSVRLTPLKIR
NVÄQYYGVSQESILGRSQSREYVLPRQVAMYFCRQKLSLSYVRIGDVFSRDHSTVISSIR
LEGGKIEENSHDIHMAIQDISKNLNSLHKSLEFFFSEEMII

CETTO 310

353472
351049

GOIN-60kDa Inner membrane Protein
YFDELSLIFRVYOMMKRTLLFVSLIGIAFVGCQIFFGYNEFRSCKNLAEKQRKISEQTLA
AVESYGLSVASMOTDTVNGEEHKNNYAVRVGDKLFLLHNGEAAQSVVSSGESMSFVDHKCG
FDNYEHLALYRQQGSSPNDTNTGKVFLPTNHEGLPVLVVEFRNNKEPLVFLGEYAQGRISN
KDSHIFGTALVFWRSGSDYIPLGLYDSREEKLVSLDLPITRAVIFONDQDSAKSSDTANN
YVITNDYMQIIVSEESGSIEGINLFFASTNNKSIVNEIGFDRDLASEKSPEALFPGLSSK
LDPGQQAKNSIGGYYPLLRRGLLSDSKKLLPLEYHALNVVSGRELATPVALRYRVLSYTP
HSIQLESLDRSVQKVYKLPENPEEKPYVFETAITLTKETEDWVTSGVPEVEIMSNASAP
TIGERVIKKNNGSLDKVKLPRVKEPLAIRRGVYPQWILNSNCYFGIILTPLSEIASGYGS
LYTGSGTAPTRLSAISPRNOLYPVSKYPGYFETLIPLP.PKDGATHRFLVVAGDLAEPTLKVL
DKB#TQEKGENPEYLDSISFRGVFAFITAPFAALLFIIMKFFKLVTGSWGISIILLTVFL
KLLLYPLNAWSIRSMRRMQILSPYLQGIOKYKNEPKRAQMEIMGLYKTNKNNPITGCLP
LLIGEPPLIAMFDLLKSSFLLRGASFIFGWIDNTAPDVLFSWGYSIWFIGNEFHTLIPL
LGIVMFLQOKVTSLHKKGPVTDQOKQQOVMCNMMAILFTAMFYNFPSGLNIYWLSSMILG
VVQWITNKILDSKHLKNEVVLNNKKHR

CPA 35311 354453 353575

CT3 1 hypothetical protein
DMARMAVIYWDRSKIYWSFEPWSLRITWYGVFFTVGIFLACLSARYLALSYYGLKDHL
FSKSQLRAVLENFFIYSILFIVPGARLAYVIFYGWSFYLQHPEEIIQIWHGGLSSHGGV
GFLLWAAIFSWIYKKKISKLTFLFLTDLCGSVFGIAAFFIRLGMFWNGEIVGTFTSLPA
VVFSDPWGGVGGVPVHPVQLYEGISYLWVSGILYFLSYKRYLHLGKGYVTSIACISVAF RFFAEYVKSHOGKVLAEDCLLTIGQILSIPLFLFGVALLIICSLKARRHRSHI

354518 354976 CPH_0312 354518 354976
CT101 hypothetical protein
CTMARNIKYFLILFROILWISAGMKLLLKATAIALDPLSSFFTYCLLSMVSWGLASLKHR
YLLSKTIRKQLSLSSEFFSQKITWIAYIKQTFISRRFLIMVIMIAFSLVLRRYINPQAL
FVIRATVGYALIKTAIAYFSKLQNALMENPEGN

acpS-Acyl-carrier Protein Synthase WKILKEISANSMEIHIGTDIIEISRIREAIATHGNRLLNRIFTEAECKYCUEKTDPIPS FACRFAGKEAVAKALGTGIGSVVAWKDIEVFKVSHGPEVLLPSHVYAKIGIEKVILSISH CKEYATATAIALA

CPn_0314 356285 355353

ETXB-Thioredoxin Reductase
MIHSRLITIGSGPSCYTAAITASRALLHPLLFEGFFSGISCGQLMTTTEVENFPGFPEGI
LGPKLMNNMKEQAVRFGTKTLAQDIISVDFSVRPFILKSKEETYSCDACIIATGASAKRL
EIGAGNDEFWOKOVTACAVCDGASPIFKNKOLYVIGYSDSALEEALTITRYGSHVYVWH
RRDKLKAKKAMEARAQNNEKITFLWNSEIVKISGDSIVRSVDIKNVTQEITTREAAGVF
PAIGHKFNTDFLXSQLTLDESGYIVTEKGTSKTSVPGVFAAGDVQVKYYRQAVTSAGSCC
LVALDARRFEG

CIm_0315 356977 rol St Ribosomal Protein 358716

1

THE BUILDING THE PROBLEM OF THE TOTAL BUILDING THE PROBLEM OF THE

KYPPGKRVLGKIVKLLPYC PEG IEGLIHISEMSWYKNIVDPSEVVNKGDEVEAIV
LSIQKDECKISLGLKOTERI EERYPIGLHYMAEIKNLTNYGAFVELEPGIBGLIH
ISOMSWIKKVSHPSELFKKGKSVEAVILGVOKESKKITLGVKQLSSNPANEIEAMFPAGT
VISGOVVIKITAFGAFVELQNGIEBLIHVSELSDKPFAKIEDIISIGENVSAKVIKLDPDH
KKVSLSVKEYLADNAYDQDSRTBLDFKDSQGPKERKKKGK

CPN_0316 359784 360121

NUSA-N Utilization Protein A

LETERLEYED IMMENICATE PROFESSIONALE GALETAAKYTLEDDANISTN

LUSTERLEYED IMMENICATE PROFESSIONALE GALETAAKYTLEDDANISTN

LUSTERLEYED IMMENICATE PROFESSIONALE GALETAAKYTLEDDANISTN

LUSTERLEYED IMMENICATE PROFESSIONALE GALETAAKYTLEDDANISTN

KTERHKIGDKIYALLYEVGESENGGAEVILSRSHAEFVKOLFIQEVPELEEGSVEIVKIA

REAGYRTKLAVRSSDENTDEVGAFVGMRGSRVKNIIRELNDEKIDIVNYSPVSTELINL

LYPTEIGKIALLEDDKVJAIVVAIVVANDAVJATVIGKGINAALISHILDYELEVQRMSETNKL

LEIQRLQLAEFDSPHLDEPLEMEGISKLVIQNLEHAGYDTIRRVLLASANDLASVPGISL

ELAYKILEQVSKYGESSYDEKPEIED

360045 362750

CPn_0317 360045 362750
infB-Initiation Factor-2
SLLIRSLSKSANMEWKLTKNLKLKIKNAQLTKAAGLDKLKQKLAQAGSSEAKSSSEKPS
AKEKSVKVALAATSTPTASAQASPESTSRRIRAKNRSSFSSSEEESSAHIPVDTSEPAP
VSIADPEPELEVVDEVCDESPEVHPVAEVLPEQPVLPETPPOEKELEPKPVKPAEPKSVV
MIKSKREPPTGKHINHLLAKTFKAPAKEEKVVAGSKSTKPVASDKTGKPGTSEGGEQNNRE
KQFNPANRSPASPKRDAGKKNLTDFRDRSKKSDESLKAFTGRDRYGLHEGGEEDRWRKK
RVYKPKKHYDEZSIGRPTHIKISLPITVADLAAEMKLKASEVIQKLFIHGMTYVVNDILD
SETAVOFIGLEFGCTIDIDYSEQDKLCLSNDTVRDEIQSTDPSKLVIRSPIVAFMGHVDH
GKTTLIDSLRKSNVAATEAGAITOHMGAFCCSTPVGDITILDTPGHEAFSAMRARGAEVC
DIVVLVVAGDEGIKEOTLEATEHAKAADIAIVVAINKCDKPNFNSSTIVRQLSEINLLPE
AWGGSTVTVTTSAKTGEGLSELLEMLALQAEVLELKADPSARARGLVIESELHKGLGPVA
TVLIQNGSLKIGEALVFNDCYGKVKTMINENNELMKEAGPSIPVLITGLSDIPKAGDPFF
VVKNEKTARDIIEARSAGOQRFALQOKKPPNFDSMLONKKTLKLMIKADVGGSIEALVSS
ISKIKSEKVDVEILTNSVGEISESDIRLAAASKAVLIGFHTGIESHAEPLIKSLGVVVEL
FTVIYHALDAIKEIMTSLLDPIAEEKDEGSAEIKEIFRSSOVGSIYGCIVTEGIMTRNHK FTVIYHA(DAIKEIMTSLLDPIAEEKDEGSAEIKEIFRSSOVGSIYGCIVTEGIMTRNHK VRVLRNYEILWKGTLSSLKRVKEDVKEVRKGLECGILLEGYQQAQIGDVLQCYEVIYHPQ

CPn_Q118 362704 363126
rbfa/Ribosome Binding Factor A
vMsynvMxLsiihkaynLkytottenRrikkrvnalloeaiakvilkdvkhpkisnLwitvt
Rvs_Skolhsarvyvsymphentkeealealkvsagfiahrasknvvlkyfpelhfyldd
ifs/PodyienLlwQioekeks

C/L_0319 363133 363879
truB-tRNA Pseudouridine Synthase
TIFFGMLMTIKDMTMDLAVELKEGILLVDKPGGRTSFSLIRALTKLIGVKKIGHAGTLDP
FATGVMVMLIGRKFTRLSDILLFEDKEYEAIAHLGTTTDSYDCDGKVVGRSKKIPSLEEV
LSAAEYFGGEIQQLPPMFSAKKVQGKKLYEYARKGLSIERHHSTVQVMLQITKYEYPLLH
FVVSCSKGTYIRSIAHELGTMLGCGAYLEQLRRLRSGRFSIDECIDGNLLDHPDFDISPY LRDAHGNSI.

CPn_0320 363824 364783

TTPISIFLPTYEMPMEIAYSLTSSFSVDSVTVGFFDGCHLGHSNLLSILTSYSGSSGVIT
FDSHPQTVLSLNHTKLINTKEERLQLLQTFPIDMLGVLTFDLNFANGSAEEFLTLLHRNL
KCKRLILGYDSCIGKEQOSNTEALDTIGKPLGIEVIKIPPYRMDNIVVSKAIRQFLSAG
NLECAHRILGHPYAISCKITEGSGIGGSLGFATINLPREESLIPLGVYACEIRYDSTTCQ
GVMNLGTAPTFGRESLYAEAHIFSFAENLYGKEVSIIPRKFLREEKKFQSKETLIRAIEK DILDAODWFAKGSFNYFGTA

365900 364767

ychF-GTP Binding Protein YSKKHVIIFIFRCLMSHTECGIVGLPNVGKSGLFNALTGAQVASCNYPFCTIDPNVGIVP VIDERLEALAKISNOKIIYADMKEVDIAGLVKKASDGAGIGNFLISHIRETHAIAHVIVP
VIDERLEALAKISNOKIIYADMKEVDIAGLVKKASDGAGIGNFLISHIRETHAIAHVIVF
CFDDPDVTHVSGKVNPVEDIEVINLELIFSDFSSAKNIHSKLEKLAKGKREVGALLPLFD
TIIAHLEKGLPLRTLELTPEQIVALKPYPFLTMKPMFYIANVDESSLPOMDNDVVAAVRE
VAAKENSKVVPICVRIEEEIVSLPIEERLEFLMSLGLEKSGLHRLVRAAVDTLGLISYT
TGPOGSRAWTVVRGSSAWEAAGEIHTDIQKGFIRAEVITFEDMIECQGRAAARELGKLHI EGRDY I VODGDTMLFLHN

CPn_0322 366231 367328
yscU-YopS Translocation Protein U
SNLGNSMGEKTEKATPKRLRDARKKGOVAKSODFPSAVTFIVSMFTAFSLSTFFFKHLGG
FLVSMLSQAPTRHDPVITLFYLKNCLMLILTASLPLLGAVAVVGVIVGFLIVGFTFSTEV
FKPDIKKFNPIENIKQKFKIKTLIELIKSILKIFGAALILYITLKSKVSLIIETAGVSPI ITAQIFKEIFYKAVTSIGIFFLIVAILDLVYQRKNFAKELKNEKFEVKQEFKDTEGNPEI KGRRRQIAQEIAYEDSSSQVKHASTVVSNPKDIAVAIGYMPEKYKAPWIIAMGINLRAKR ILDEAEKYGIPIMRNVPLAHQLLDEGKELKFIPESTYEAIGEILLYITSLNAQNPNNKNT NOPDHL

367322 369460

CPn_0323 367322 369460
lcrD- Low Calcium Response D
59FIMNKLINFVSRTLOGDTALMMINKSSDLILALMMMGVVLMIIIPLPPPIVDLMITINL
SISVFLLMVALYIPSALOLSVPPSLLLITTMFRLGINISSSRQILLKAYAGHVIQAFGDF
VVCGNYVVGFIIFLIITIIOFIVVTKGAERVAEVAARFRLDAMPGKQMAIDADLRAGMID
ATQARDKRAQIQKESELVGAMDGAMKFIKGDVIAGIVISLINIVGGLTIGVAHHGMDLA
AAHVYTLLSICDGLVSQIPSLLIALTAGIVTTRVSSDKNTNLGKEISTQLVKEPRALLLA
GAATLGVGFFKGFPLWFSFILLALIFVALGILLITKKSAGKKGGGSGATTVGAADGAA
TVGONPDDYSLTLPVILELGKDLSKLIGHKTKSGQSFVDDMIPKMRQALYQDIGIRYPGI
HVRTDSPSLEGYDYMILLNEVPYVRGKIPPHHVLTNEVEDNLSRYNLPFITYKNAAGLPS
AWV:EDAKAILEKAAIKWTPLEVTILHLSYFFHKGSQEFLGIQEVRSMIEPMERSPPDL
VKEVTRLIPLQKLTEIFKRLVQEDISIKDLRTILESLSEWAQTEKDTVLLTETVRSSLKL
YLGFKFSGQGSAISVYLLDFEIEEMIRGAIKGTGMISYLALDPDSVNLILKSMRNTITPT
PAXXXPPVLLTAIDVRRYVRKLIETEFPDIAVIGYQEILPEIRIQPLGRIQIF

CPn: 9:24 10:1468 370688

CTH_0_024 Hopotherical profesion
YVVAIIRRIMAASCATGGLGOTO/MILAAVEAAAKADAAEVVAQOEGSEMNMIQOGQDLT
YVVAIIRRIMAASCATGGLGOTO/MILAAVEAAAKADAAEVVAQOEGSEMNMIQOGQDLT
NEAAATRIKKKEEKRÖTLE:RRKKGEAGKAERYDESTEEKREDDILADKYA:GOJSE LGODEL
RGLEDDI GIDDAGTEDI CLALVQEK IKBIJALOJTAALDYLVATTPPBQGKLKEAL IQARNTHT
EQFFIRTA IGAKNILFAJOEYADGLIVOERGLEOD/ILEVTIJTHITGOOLISMLQDRYTYYQD
MA IVGGFLMKGMATELKRQJPY/PGAQLJVUJATETRII/JAVETGYDYDYFESRVPILLDGLK

AEGIOTO:DUNFVKVAEGYHKIINDKFPTAGKVEREVI OTGGRUFGGADKROQU/JAMIANALDAVNINNEDYPKAS

VDSVTGVLNLFFSALR

CPn_0325 370596 371148 CT325 hypothetical protein KRIAMONOYEQLLESLAPLLNTTLAPDKNNSCLIRFSDTHVPVQIEEDGNSGDLAVSTLL GTLPENVFRERIFKAALSVWGSFQSSIKGILGYGEVTQQLYLSDILSMNYLNGEKLFEYL KLFSLHAKIWMESLRTGNLPDLHYLGIYYVA

CPn_0327 372927 373211 r128-L28 Ribosomal Protein RIHRKUMSRKCPLTGKRPRRGYSYTLRGIAKKKKGIGLKVTGKTKRRFFPNMLTKRLWST EENRFLKLKISASALRHIDKLGLEKVLERAKSKNF

CPn_0328 373220 374992
CT085 hypothetical protein
LKYREIFMSFLRRHISLFRSGVGLIDVFAPVSPNLELAEIHRRVIEDGGPALLFHNVIGS
SFPUT.HLFGTKHRVJOLFSOAPDNLIARVAHLISSTFKLSSLMKSRDLLKRISSLGLKK
ARFRRFPFVSMSSVNLDHLPLLTSWPEDGGAFLTLPLVYTESPTLTTFNLGMYRVQRFNQ
NTMGLHFQIOKGGGMLYFAAEOKKQNLPVSVFLSGNPFLTLSAIAPLEPNNSELLFATFL
GGAKLLKKKTNDHPHPLLYDAEFILVGESPAGKRAPEDFPGDHFGYYSLQHDFPEFHCHK
IYHRKDAIYPATVVGKPYQEDFYIGNKLQEYLSPLFPLVMFQVRLKSYGESGFHALTAA
VVKERYWRESLTTALRILGEGQLSLTKFLMYTDGDPVLDRFSVVLETILERLQPDRDLIT
FSETANDTLDYTGPSLNKGSKGIFMSIGKAIRDLPHGYQGKHGVQDIAPFCRGCLVLE
TSLEDRGIKSLLHHPDLKSWPLIILADNLRETIQSEKDFLWRTFTRCAPANDLHALHSHF
ATHRPNYNFPFVIDALMKPSYPKEVEVDPSTKQKVSERWHAYFPNKETFYI

CPR-0329 375085 376146
Phopholipase D Superfamily (leader (33) peptide)
Phopholipase D Superfamily (leader (33) peptide)
MMRKROKDKLKICVIISTILLVGIFARAPRGDTFKTFLKSEEAITYSNOCHEDMRKILCD
ALEHADEE IFLTYNLSEPKIQOSLTROQAGAKKVTITYYOKK KIQILKQASNVTLVEOP
PAGRKLMHOKALSIDKKDAWLGSANYTNLSLRLDNNLILGHISSELCDLITTNTSGDFSI
KDÓTGKYFVLPODRKTAIQAVLEKIQTAGVATIQVAMFALTHSEIIQALHQAKORGIHVDI
ITBRSHSKLTFKOLROLINKDFVSINTAPCTLHHKFAVIDNKTLLAGSINWSKGRFSLN
DZSIILLENLTKQQNQKLRMIWKDLAKHSEHPTVDDEEKEIIEKSLPVEEQEAA

252.0330 376930 376202
CÉN_0330 376930 376202
CÉN_33 hypothetical protein
FISIEMLLISROLFSVLPSRFODLHVYRFKESLKLLOFMTMVGGEIVVVLAEIKEEDLR
RÉÉVRKRREKNYLRIFRVLSRFDWRIIRFDPYGALSAQSIAKDSRONSPLVEKISEEN
ATNEAIRIALLAIGDREQEEKKORHRYKLLGOKQAKVLLSQLRHVHLDFKKLYCDSKKKA
DÖEKDEKNKOKRSIKVTKKKKGISLGAAASQAIAAAAEAWVIARNKGVLETASTLFYQKD

CER_0331 378452 376701

CER_0331 378452 376701

COUNTY OF THE COUNTY ON THE COUNTY OF
378676 378536 CPn_0332 CHLTR T2 Protein

YLDSRIRVIPLARORCTLLHLLAVLCPPISFFTQGVSPCVFFCFLD

378800 379117 CPn _0333

TELB VDEFVFFMGKPKKSRTDRALAGEIQKKSTEVLKKPARIKANGRRKFLIAKEOKTLKHR AGEYDQLVRSLLDSQKKDTDKVLIFNYENGFVFTDKDHFSKY9IRL

379308 379823

CPn_0334 379308 379823
CTOT9 similarity
TMSVHITPRKCFILCILSMFTLPTLFPKAHLILFSPYIVYCFYCFSKDKGLVLALGCGVL
SDLALGSGGVFLLLYPLTALITHKAHLIFSKESKAALVYVMIFYGVFLLLTIPMCALFG
HEVRWSIDVLMIPLKCSFLDNLIFTSVIYILPCAINSGTHKMISFFRRLVCY

CPn_0335 379808 380674

to/ID-Methylene Tetrahydrofolate Denyirogenase

EIGMLLPG/FPAAEKILDRUKKEFISOSPTSPGLAVV/LIGNDPASEVYVGMKVKKATEIGII

SKAHKLPGDSTLSCVLKLIERLNODPSIHGILVOYPLPKHLDSEVILDAISPDKDVGLHPVMMCKLLLONFOGLLPGTPAGIIELLNYYEIPYRGRHAAIVGRSNIVGKPLAALMMOKHHYDNCTYVUHCSSENLPEILKTADIIIAALSPEIRETMYAPHAVIVDVGTTRVPADNAKGYTLL/JDVDFNNV/TKGAAITPVPGGVGPM/VAMLMSNTWRCYONFS

380569

70 jl.
RODKMEDBIBGDDWRRWGILLGRYAHEQYMANLEKFFLVLLGLGLGDGGKTTT LEGEOMFL
FYR FYLTYTGGGAKEKAGLGGG LDRGEHK DIG LYNNWNFYSELG LLNRAPADVPLTLGWEL
GEFLDOVIZHJYKLGEGRFOLTVOPLKTIWILLHLKGGTLEPKDVWGGHYKDMGWGHLEFGG
HYRFLL FRANFING LDLAGVYKNY AVDOLME LONTEG PRNOVUMS KEI EKTSGHIBGGRWR
LFGRAAGT LLOT DI MA LATGONI LOKWCVFGK LYTH LLDRTGKFLELSSYP LOSVSVVH
RGGAYADA LATVLMTETGK LEAKOWEEHH LLTY LIDGAGS

CPT_U33;
smpB- Small Protein
leEIFPGNOGKRILIIVLRRKNCFLLYWFLGPIMGEDLMAGKEIVSNRKALRNYEVIETL
LEAGIVLTOTEIKSLRDHOGNLGDAYV.VSKGEGWLLNASIAPYRFGNIYNHEERRKRKLL
LHRYELRKLEGKIAQKGMTLIPLGMFLSRGYVKVRLGCCRGKKAYDKRTIIEREKEREV AAAMKRRHH

383375 ·in

39227 PONTO LEVILIBRAL (ETVIDEL VETATOLIVATEGYTK ENMERT GRNES MILLER TO A FLATER THE LEFT THE LART HER AFTER TATE OF A RAYYERGA IS LEGGREFIC WELL FRANCE IS SAGEMA, IT JOSS CERLISHERE DEPMI PDI ONALIFS LIPAGULKTHE RESPANSEES RYVITOYLLA IANGVAT IVOTOGRILAK IDAEVILDKEFSGEVI IP I MAVEET I KNCSDEGEAA I FLODK I AVECDITLLITKLISG EFPDFSPVISTESNVKLDI MREEL I TILLKOVALFTNESSHSVKFSFLIGELTLTANCTKV GECKVSMAVNYSGELLE JAFNPFFFLDILKHSKDELVSLGISDSYNPGI I TDSASGLFVI MANDELHID.

383405 384034

CPn_0339 383405 384034
CT339 hypothetical protein
VTTLPMFMKICSLKLMFRNHSDLEISLAPKLNYAQGKTNLLEALYVLSLGRSFRTOHLT
VTTLTFGSSHFFLETOFEKOHLEOALSIYTDKQGKKICYNQLPIKTLSQLIGKVPIVLFSS
KDRLLISGAPADERLFINLLLSQCDNHYTLCLSYYHRALQQRNALLKSKQTSTVASGNS
WSMTAPTYPSNGFSVVRNFQIYPKNFGLTT

383842 384156 CPn 0340

CFT_0540
(Frame-shift with 0339)
PLYPLLIVISRSSAEKCSLKKQANLNRGLWDEQLVKHGTYLSIQRFLCSQKLSDLSKEL
WSNNLKEQLALKFKSSLIKNSDISETAVAEEFHKQLSISLPRDLE

91 384160 384495 Shift with 0340)

(frame-shift with 0340)
GSTSVEPHREDFLLTMNOMFVSQFSSEGQKHSLLAILRLAECLYLKQSHHVSPLVCLDDI
HAGLMAERVGQLLDPAPTLGQTLITSTHMHGELPKTSLVLSIENAQVSEQII

CPT_0342 384619 385062
predicted OMP [leader (19) peptide]
HMKKFLLTILFLAVCNPLFSETSVIOTLPSGIGGLKETSKOKESVVCVHAFLRSYTSLKP
FARVLEKEHYDVF IMNYETRKFTLEKHAEHLNRLLKKIAELKPGVPINFVTHSIGGVIVR VALAHPDCPEEAKKEKLFSWLLRTQGLH

385595 384999

CPn_0343 J84999 J85595

(frame-shift with 0342?)

LPRRSQKRKAILMAPPNAGSTLARRYRCVKFVQFVFGGKLGRQLLTYCPTKMLNVGKLPS

SLDVLILSGNRHSKFLPFRLPYENDGKVCTIETKLDTPHKAYVIHTSHTYIITNKSLYL

MKEFLKEGNTTPIIEHVPEAALEQTVMEDKQKNSRLKPYPNQDIYVIHCFGSRPYNLYGF PKKWSLNQKNEINPEKLEK

CPn_0344 387432 385558

yael-Metalloprotease
SSRWMTIYFILAALALGILVLIHELGHLVVAKAVGMAVESFSIGFGPALFKKRIGGIEY
RIGCIPFGYVRIRGMERTKEKGEKGKIDSVYDIPQGFFSKSPWKRILVLVAGPLANILL
AVLAFSILYMNGGRSKNYSDCSKVVGWVHPVLQAEGLLPGDEILTCNGKPVVGDKDHLTT
SLLEGHLMLEIKRFGYLTVPSKEFAIDVEFDPTKFGVPCSGASYLLYSNQVPLTKNSPME
NSELRPNDRFVWMDGTLLFSMAQISGILMESYAFVAVARNOKIFFSRQPRVLASVLHYTN
YLNNELIDTQYEAGLKGKWSSLYTLPYVINSYGYIEGELTAIDPESPLPOPQERLOLGDR
ILAIDGTPVSGSVDILRLVQNNRVSIIVQQWSPQELEEVNSRDADKRFIASYHSEDLLQI
LIHHLGESHPVEVAGPYRLDPVQPRPWIDVYSSESLDKQLEVAKKINKNKDRYYLERLD
AEKQKPSLGISLKDLKVRYNPSPVVMLSNITKESLITLKALVTGHLSPQWLSGPVGIVQV
LHTGWSVGFSEVLFWIGLISMNLAVLNLLPIPVLDGGYILLCLWEIVKRRRLNMKIVERI
LVPFTFLLIHIFIFLTFQDLFRFFG LVPFTFLLIIFFIFLTFQDLFRFFG

388587 387436

CPn_0345 388587 387436
CT345 hypothetical protein
LKVACLKHLAVLGSTGSIGROTLEIVRRYPSEFKIISMASYGNNLRLFFQQLEEFAPLAA
AVYNEEVYNEACORFPHMOFFLQGEGLTQLCIMDTVTTVVAASSGIEALPAILESMKKGK
ALALANKEILVCAGELVSKTAKENSIKVLPIDSEKNALYOCLEGRTIEGIKKLILTASGG
PLINKSLEELSCYTKGDVLNHPIMMMGSKVTVDSSTLVNKGLEIIEAYMLFGLENVEILA
VIHPOSLIHGMVEFLGGSVISIMNPPDMLFPIQYALTAPERFASPRDGMDFSKKQTLEFF
PVDEERFPSIRLAQQVLEKQGSSGSFFNAANEVLVRRFLCEEISWCDILRKLTTLMECHK
VYACHSLEDILEVDGEARALAQEI

CPn_0346 389690 388704
070-troD/ytgD-Integral Membrane Protein
KKGSIMALGPSPYYGVSFFQFFSVFFSRLFSGSLFTGSLYIDDIQIIVFLAISCSGAFAG
FELVLRMAMYANAVSHTVLFGLVCVCLFTHQLTTLSLGTLTLAAMATAMLTGFLIYFIR
NTFKVSEESSTALVFSLLFSLSVLLVFMTKNAHIGTELVLGNADSLTKEDIFFVTIVIL
ANAVITIFAFRSLVCSSFDSVFASSLGIPIRLVDYLIIFQLSACLVGAFKAVGLMALAF
LIIPSLIARVIAKSIRSLMAWSLVFSIGTAFLAPASSRAILSAYDLGLSTSGISVVFLTM
MYIVVKFISYFRGYFSKNFEKISEKSSQY

391078 389678

CPn_0347 391078 389678
069-troC/ytgC-integral Membrane Protein
TFGTNPEALSRKTIWIVLIMLSCVFSDTIFLSSFLAVTLICMTTALWGTILLISKQPLLS
ESLSHASYPGLLVGALMACYVFSLQASIFWIVLFGCAASVFGYGIIVFLGKVCKLHKDSA
LCFVLVVFFAIGVILASYVKESSPTLYNRINAYLYGQAATLGFLEATLAAIVFCASLFAL
WMWYRQIVVTTFDKDFAVTCGLKTVLYEALSLIFISLVIVSGVRSVGIVLISAMFVAPSL
GARQLSDRLSTLLILSAFFGGISGALGSYISVAFTCRAIIGQQAVPVTLPTGPLVVICAG
LLAGLCLLFSPKSGWVIRFVRRHFSFSKDGHLLKVFWHISHNRLENISVRDFVCSYKY
QEYFGPKPFPRWRVQILENRGYYKKEODYYRLTKKGRSEALRLVRAHRLWESYLVNSLDF
SVECVHELABELEHVLTEELDITTLEFILNDFCYDDHROITFDKKKFV SKESVHELAEETEHVLTEELDHTLTETLNDPCYDPHPQTTPNKKKEV

CPn_0348 391815 391027

068-trob/ytgB-ABC transporter ATPASE
FCWLLNYPDETFWSVNNLCVNTEHAAVLYHIGFGLGKGGLTAILGPNGAGKGTLLKASLG
LIKPSGGTTYFFNGKFKKYRQFIAYMPGRACYIMDFPMTVLDLALMXYGYKGMWGRISS
DGRREAFHLERWGLEGVAPROIFGGLXXGQGPAFLAKALMGKADLYLMDELF:AIDMAS
FKTSVGVLQELRQGKTIVVVHHDLSHVRQLFDHVVLLNKRLIGGGPTDECLNGDTIFQT
YGGEIELLEOTLKLGRGKGFGGT

30170 cpn_0349 - 予覧権7 391790 - 967-troA/yrgA-Solute Protein Binding Family WILKNA:REMDAKNCYIFKVMRWIFCFVACGITFGCTN
VERVVCHRLATAVLIKGSLDPHAYEMVKGDKDKIAGSA
PHJYKLGERLIARGAFVPLEEDGICDPHIWMDLSIWKEA
TEVLIEKFPENSAEFKA
NSEELVCEMSILDGWAKGCLSTIPENLRYLVGSHANAFSYFTRYLATFEEVASGAWRSRC
ISPEGLSPEAQISVRDIMAVVDYINEHDVSVVFPEDTLNQDALKKIVSSLKKSHLVRLAQ
KPLYSDNVDDNYFSTFKHNVCLITEELGGVALECQR

CPn_0350 393169 393684

THE STATE OF THE S ARFSPKKPLTSLKRELIRSIRNGIVSVELWNAYVEAVKAVSSPNLEVTSPFV

393861 395432

CPn_0351 393861 395432
adt-ADP/ATP Translocase
kikvforvnntkteekpfcklrsflwpihthelkkvlpmflmffcitfnytvlrdtkdtl
kikvforvnntkteekpfcklrsflwpihthelkkvlpmflmffcitfnytvlrdtkdtl
kikvforvnntkteekpfcklrsflwpihthelkkvlpmflmffcitfnytvlrdtkdtl
kikvforvntpteffackoailfpclglvailrnwffaaptvlaelkosvnuslimfwaf
kivflkheakreyalfgiganisllasgraivwasklrasvsetvdpwgislrllmamt
kivsglvlmasywminknvltdprpfynpeemokokkgakpkmmkosflylarspyillla
klviavgicinlievtwksolkloypnmndyseffwdrsfymgvvsvlihltvggnvirk
fgwltgalvtpvwylltgivffalviffnoasglvamfgttplmlavvvgaionilskst
kyalfdstkemayipldoeokvkgkaaidvvaarfgksggalioogllvicgsigamtpy
lavillfiiaiwlvsatklnklflagsalkeoevagedsapass

CPn_0352 395478 396830

No robust homolog present in Genebank/EMBL as of 11/7/98

WVGIFFINSHFTNSYAFFNQKVIITVRHSGCTMKCSPLTLVPHIFLKNDCECHRSCSLKI
RTIARLILGLVLALVSALSFVFLAAPISYAIGGTLALAAIVILIITLVVALLAKSKVLPI
PNELOKIIYNRYPKEVFFVYTHSTITVNELKIFINCWRSGTDLPPNLHKKAEAFGIDILK
SIDLTLFPEFEEILLQNCPLYWLSHFIDKTESVAGEIGLMKTQKVVGLLGPLAFHKGYTT
IFHSYTRFLLTLISESQYKFLYSKASKNQWDSPSVKKTCEEIFKELPHMIFFKDVQGIS
OFLFLFFSHGITWEQAQMIQLINPDMWKHLCOFDKAGGHCSMATFGGFLMTETNMFDPVS
SNYEPTVNFMTWKELKVLLEKVKESFMHPASALVQKICVNTTHHQNLLKRWQFVRNTSSQ
WTSSLPQYAFHAQTYKLEKKIESSLPIRSSL

CPn_0353 396893 397135
No robust homolog present in Genebank/EMBL as of 11/7/98
LRFRNIKKSLIFIKRIRYSOSGKEOKGARPFFKKSITSSLVILLLEAIFNENFSSIIONN
FNKNFKNKNISINRIFVKFTI

CPHEO354 397062 398507

NO FODUST HOMOLOG PRESENT IN GENEBARK/EMBL AS OF 11/7/98

YKTESIKILKIKHTFLLIGFLIALRYNTOIDEPRKCMSNITSFVIQNORSCNYYFELNNST
TIHTVISAILLCGALTAFLCVAAPVSYILSGALLGLLIALIGVILGIKKITPMISSKE
OVFFDELVNARIRAHYRFYSDFVSEAKFNLKDLISFIDLIAQLHSEVGSSTNYNVSEELQ
OKEDTFEGIARLKNEVRTASLKRLESAASSRPLFPSLPKILOKVFPFFWLGEFISAGSKV
VELERVKKIGGSLEEDLSSYIRPENLPTYWLIPLDFRPTNSSILNLHTLVLARVLTRUFF
OHLEYAALNGEWNLINSDLINTKKOQLFAKYHAAYQSYKHLSQPSLQEDEFYNLLLCIFKH
RYFWKQMSLIKTVPADLWENLCCLTLDHTGRPQDMEFASLIGTLYTQGLIHKESEAFLSS
LTILESLDQFKTIRRQSTNIAMFLENLATHNSTFRSLPPITVHPLKRSVFSQPEEDESSLL

CPHE0355 399955 398591

NOTED BUT TO THE PROPERTY OF THE PROPE HRIALEHGNARARVYDVNFVTGARIHRKTSIFFKD

CPT=0356 400465 400109
No Edbust homolog present in Genebank/EMBL as of 11/7/98
KQVGLFQYMNESGWDWLCDFDSQGGGFQLSRLVGLLHSSWALYEAKEGFYLPEVSLI ELTEMOLLSKPTKHGVAKDLCNVFEKHFQRFRQYLGSLDLNQRFENTFLNYPKYHLDF

401341 400469 CPN_0357 401341 400469
No robust homolog present in Genebank/EMBL as of 11/7/98
YSSHNGASMYNIOPVYRNTOVNYSOATOFSVCOPALSLIIVSVVAAVLAIVALVCSOSLL
SIELGTALVLVSLILFASAMFMIYKMROEPKELLIPKKIMELIOEHYPSIVVDFIRODEV
SIYEIHHLISILMKTNYFOKAPVYLOEKILOFGEIKFKDVHFESKLPNFEELLLOFCPLHW
LGRLVYPMVSDVTPCTYGYYWCGPLGLYENAPSLFERRSLLLLKKISFGEFALLEDGLKK NTWSSSELVQIRQNLFTRYYADKEEVDEAELNADYEQFDSLLHLIFSHKLS

401757 401578 No robust homolog present in Genebank/EMBL as of 11/7/98
EEVLSVOMKLIPTODSIERETDSKRDKKIFTIYICSSKVLAGHFFSHLDKHNKIHESIGV

401994 403817

lepA-GTPase TITLOYILKEYKIENIRNFSIIAHIDHGKSTIADRLLESTSTVEEREMREQLLDSMDLERE RGITIKAHPVTMTYLYEGEVYQLNLIDTPGHVDFSYEVSRSLSACEGALLIVDAAQGVQA QSLANV/LALERDLEIIPVLNKIDLPAADPVRIAQQIEDYIGLDTTN/IACSAKTGQGIP QSLANVILALEROLETTPVLNKTDLPAADPVRTAQQTEDYTGLDTTNITACSAKTGQTP
ALLKALIDLVPPPKAPAETELKALVFDSHYDPVVGIMVYVRTISGELKKGDRITFMAAKG
SSFEVLGJGAFLPKAFFIEGSLRFQQVGFFIANLKKVKDVKIGDTVFKTKHPAKTPLEGF
KEINPVVFAGIYPIDSSDFDTLKDALGRLQLNDSALTTEQESSHSVGFGFRGGFLGLHL
ELTFEPLIREFDLDITATAPSVIYKVVLKNGKVLDIDNPSGYPDPAITEHVEEPWHVNI
TIFKDEYLSNIMNLCLDKRGICVKTENLDQHRLVLAYELDFLNETVBDFNDKLKSVTKGYGS
FDYRLGDYRKGSTIKLEVLINEEPIDAFSCLVHRDKAESRGRSYCEKLVDVIPQQLFKIP
190ATHIKKVTARETTRALSKNYTAKCYGGDITRKRKLWEKQKKKKKKMKEFGKVSIPNTA FIEATKID

403922 Ctm 0360 405364 CTI. 0.3.0 AD3.00 AD3.00 AD3.00 CTI. 0.0 AD3.00 AD3 ALKOPEGIPVIDTILDVVOG HALIDALNSOVPLOLLIGAVLARFLOSWKEIREÇA ARMYPOTPLIFEMPHOPSVF HALIVASKIISYAGGFMLLJEASKEYMMGLDLGEIA LMWRGGCIIQSAFLDVIHKGFAANPENTSLIFQEYFRGALRHAEMGWRRTVATAIGAGLP IPCLAAAITFYDT/PTASSSMSLAGGLRDYFGAHTYERNDRPRGEFYHTDWVHTKTTERV

CPn_0361 405650 405382

EVTS-tyrosyl than synthetase
CLAIMITEDIMENT OF ACTIVITIES THE INTERPOLATION OF PETAPALHIGHNIOL
THER LAAD HITE CLAIM MALHENER MEDIAL THE WHODGO KITAGE, BYLD TO
TUVNADALQEIDLIDFULGKHIPK DAMINENTIK QRVHODGO JYTEFOYLILOSYO
TYHLEKNYGTILCCGSDOWNITSGIDFIRRKGLGQAYGLTYPLLTNAQKKIGKTESG
TYWLDSDLTSPELYQYLLALPDDTIPKIARTLILLSNEEIQDIRRVOTDEVAVXEFVA
QDILSAIHGDLGLEFALSYTRSMHERNLSSLSEKOPHELFAGGMASLDKSEVLGKRWLD
LFLVIGLCKSKGETRRLYEQKGYYINNVPIANEHSVCEEQDICYGHYVLLAQGKKRKLVL
YLN

CPn_0362 407843 407055
flia/rpsD-sigma-28/WhiG Family
LDKKKFVKTQOTQNFIEWNFYWETQEIEYRDSLIEFYLPLVKSVVHRLISGMPSHVKTE
DLYASGVEGLVRAVERYNPERSRRFEGYAVFLIKAAIIDDLRKQDWVPRSVHQKANKLSG
AMDSLRQSLGKEPYDLELCEYLNISQOELSGWFVSARPALIVSLNEEWPSQSDEGAGMAL
EERIPDERAETGYDVVDKQEFSLCLANAIQELEEKERKVMALYYYEELVLKEIGKVLGVS
ESRVSQIHSKAL/KLRAALSAFR

409700 407943

CPn_0363 409700 407943
flha-Flagellar Secretion Protein
eavrysgkkdgyrgmifylsilvliftplpoilldfglcisfalslltvcwyftlnssn
saklfppfflylclirlglnlastrwivssctasslivslgsffslgslwaatfaclllf
fvnflwvsygseriaevrsrfflealpakomaldsdlvsgrasykavkkonalieedb
fsamegvrffykgdaiiscilllvnvvsvtclyytsgyaleomyftyldablysovpall
scaaatliskidkeesllnvlffykglrohfrvvsllifslccipsspkfpivllasl
lwlayrkepasedscierafsyvegacpkedesofyovyraaseevfedlgyrlevlts
lrieerphlryfgonvyldemtpeavlpflrniahealnaev/okyleeseryfglavlts
lrieerphlryfgonvyldemtpeavlpflrniahealnaev/okyleeseryfglavgwi
grslwdokotlevitidphveelinssysksnpvwgenvirrvdsllersvfkdfraivt
scettfemkwldphfpdllvlshdelpkeipisflgivsdevlvp

CPn_0364 409954 410238
feg/4-Ferredoxin IV
KENSMAXLVITSDDEQQEFELEDNSEIAEPCESMGIPFACTEGVCGTCVIEVLEGRENLS
EFTEPEYDFLGEPEDSNERLACQCRIKGGCVKVTF

Pn_0365 410498 411544
No robust homolog present in Genebank/EMBL as of 11/7/98
FEKTQVNSLIMATISPISLTVDHPLVDTKKKSCSNFDKIQSRILLITAIFAVLVTIGTLL
IGLLLNIPVIYFLTGISFIAVVLSNFILYKRATTLLKPRACGKHKEIKPKRVSTNLQYSS
ISIAINRSKENWEHQPKDLQNLPAFSALLTDNPYEIWKAKHSLFSLVSLLPGGNPEHLLI
SASENLGKTLLIEETSQNAPISSYVDTTPSPKSLLNEAIQETRVEINTELPAGDSGERLY
WQPDFRGRVFLPQIPTTPEAIYQYYYALYVTYIQTAINTNTQIIQIPLYSLREHLYSREL
PPQSRMQQSLAMITAVKYMAELHPEYPLTIACVERSLAQLPQESIEDLS 410498

CPn_0366 411976 412440
No robust homolog present in Genebank/EMBL as of 11/7/98
MGYLPVSATUVLFESPAAPLINSANTONOKLIELKGKQQAESSPRTITSVILEVLLVIGC
CLIVLSLLAIRPALGFTLETGHPAAIAVLAVSGTILLVAVIILFCFLAAVPFAAKKTYKY VKTVDDYASWHSHQQTPTLGTIFSGIVYAESQAQL

413078 413836 CPT_0367 413078 413356

No robust homolog present in Genebank/EMBL as of 11/7/98
SFPLNRYFMTKTTSIPDVHENQSHLSVDERLISESPVLTKKEVIAKIIKLTALILAIA
VGTAVVAGVUGMPLMAIATGAALLAAVVLSCLLLRREPSKPTEELLGPQKHVPKDLAIA
VQPSVPLDYQKLLRREMTLVNTLSEINISMTLOPDQRYVWEHQGAPITLVATTGDIAK
PRIKTSGRYMIVNAANSNMQSGGAGTNAALSAATHPTCWNNTRTSGGKINTGKGLSVGEC RSAPWINRDWINK

CPn_0368 413766 414107
No robust homolog present in Genebank/EMBL as of 11/7/98
TLAKDYLWVNAAQHPGSIETGRINDTNPGEAHFLAQLLGPKYEGELKAHPEKLSNVIKKA
YLNCFDEALNNQATVVQVPLISSSIYSPGGKLELEPVNQTKPNSSAYKLYHIRT

0369 414345 415562 CPH_0369 414345 415562
CTO58 hypothetical protein_2
NIMTDSNPLPSYTCASLYRTPAKHSYPIRLPLNRTDRIEKILKIVTLTLALACALGFSIA
AGILAMPIFSAVVITLAIAAVSLYSLLKKRPKLYEILPOIEPESEQSSLSPSPOPPEQOD
LPLQIDPLPUPPESLEPSYLADLTTPPEELTAITVTPGYEALLEONWOLLPSLAAVDPSFT
TETPOQPCFIWKLKDSKLIFISTSCDIAVPRIKTGGRVMIVNAANENISREGGGTNRALS
LATSLOCWNASRLFRAHSRSGSQLOPGECRSAKWENSDHTSNDHVPGKAHFLAOLLGPEA
AKCNNDPKQAFEVSKKAFHNLFQEAEIIGVDVIQLPLIGCNLFAPSRLLNLGKTRAEWIE
AIKLALITSLQDFGWEQDNQEEQKIIILTDKDQPPIIPPRDLTTP

415755 416912 CPH_0370 415755 416912
CTO58 hypothetical protein_3
KRIFFKLFVFYLKSFMSTTEPNLTNVNLTMLISSESMPTQLASHKLKGLDLVAFILIIGI
AVSSGTAAIILGIFLLFILTALAVLAFSILLYFLLREKRSFISVTHQPTPIIKDTDLPPV
PPLALTPVPTEATLEEPPLFSPRTHOTLLQENMORIPDLQANTOMPFIAADNQTGYAMHL
KNSNLTLISTLGFIEKPRYKTGGIVMIVNAATPKNAANVKOTSLALAKATSVRCWENSKK
SPDPLRSKQPLQLJECRSAKWENLKGTTNAGKAGLPQFLQDLLGPKASDYNYNPNDAFFE
CRQAYLDKLMEAKRRKTTVACLPLLGSHFPGSPKDEETTSLRLQWIDGVKLALIDALQTF
GSEAENQNOPWVIILTTLARHPLITP

CPn_0371 417141 417503 No robust homolog present in Genebank/EMBL as of 1177798 KTMPVSSAPLPTSHRPSSONLILMERENKALKARHQDETPKTIKLLVKILVATLVIEVLG LIAAFFIRSTPPICLITICALILTTY/CVLLL/YIKLALVHKTEXTPARQQIKRKLSSKSI

418061 CPn_0372 CPR_0172
No robust homelog present in Genebank/EMBL as of LL77798
NYRACHRIGHNESSEVVTGTGGGGPVEGTFF/GEFFEREGGGGGGGTKIAFAAGTALLLIN
TEVOGIVAIAMIEVATSVGAYFTYIGFLFLLGJELGAIMLEMYKITHPSQNTFISN 418356 420218

GCPE
NSEIFEIFMTLITPAINSSRRKTHTVRIGNLYIGSDHSIKTQSMITTLITTDIDSTVEQIY
ALAEHNCDIVRVTVQGIKEAQACEKIKERLIALGLNIPLVADIHFFPQAMLVADFADKV
RINPGNYIDKRNMFKGTKIYTEASYAQSLLRLEEKFAPLVEKCKRLGKAMRIGVNHGSLS
ERIMQKYGDTIEGMVASAIEYIAVCEKLNYRDVVFSMKSSNPKIMVTAYRQLAKDLDARG ERIMOKYGOTTEGMVASALEYLAVCEKLMYRDVYSHASSNEKIAVJARQLADLLANG WLYPLHLGVYTEAGMCVDGILKRAVIGOTLAGGIGGTILAGGIGGTIRGSUTGCOPTTEIPVCOSLLANG WLYPLHLGVYTEAGHCDGILKSANGENAKTTIAKIDVSYFLKLYFHHLTOFTPEELLEH GUGGAFFEKARTTIAKIVVYHEBLKDALTDYLISEHFLYPHHIMOVIGLYFHINEELWODDALHAVYHVHASDPFHTSRDFFEKQGHQKPTKLVFSRDFDKMEEAALSIATEFGALLL DGLGEAVVLDLENLPLQDVLKTAFTTLOMAGYRLKKTEYISCPMCGRTLFDLEEVTTRIR KRTQHLPGLKIATMGCIVNGPGEMADADFGFVGSKTGMIDLYVKHTCVKAHIPMEDAEEE LIRLLOEHGVWKDPEETKLTV

CPn_0374 420209 420961
CT056 hypothetical protein
vDsmtLsfhTheLnywTfEEFFOCLPIRHGVFSKQKDAEGTVFAAKNPEIASALQSPKYCD
LHQRHGTSVRCVTPTSPTYQPADGLCTQSPLLSLHIRHSDCQAAIFYDREHHAIANVHSG
WRGLLGNIYAVTVGTMKKLFHTKPODLFVAIGPSIGPDYAIYPDYATLFPRSFLPFMNPK
NHFDLRAIARKQLTNLGISKDRIFISDLCTYTEHDAFFSSRYLAHHPDPNLTGQHSKNRN

421615 421112 CPP_U3/5
No robust homolog present in Genebank/EMBL as of 11/7/98
RLSMKLGASTNKKVHEPVKPKKAQLAEIEANKTQATECTLRSKSLALQIARAVLYILFAA:
LMLAAGITFVTFLALGFPLIQAYSIAGIITLVGLAIGVLLILSLLPKEDEEADALSRNA
LLPLTIIVTEQOPITPKPEIPYSYLTKLALLTSLFLTLRRSSSQRKTH

CPn_0376 421680 422294

No robust homolog present in Genebank/EMBL as of 11/7/98
FKVVTAKAPNLTEIRDHGARVPSLFLLSPETSHWKGDKEVSAPLKOLODLLGEEQWEANK
TKMNSRKKAGGWAIFNSPTGVSSTLVLAWTPWYDKDVODILERKDPMSSSLSEKDSK
EFLKNLFVDLLENGFTSVHIHAEEAFTPLDHTGKPHFKRDNVYLPGKLIGALNEAAVQAN VSADTOFTLFLTQDECNPFHDKKRG

CPn_0377 423441 422347
sucB-Dihydrolipoamide Succinyltransferase
IMTFEVRIPNIAESISEVTVASILVYEBGALIOENGGLLEIESDKVNQLIYAPVSGRIFWE
VSEĞDVVYGGVVGKI EPAGEGEELGDSGSKETIEAEIICFFQSGVRQSPPENKTFIPLR
DOMDOGSGGLSAGDRGETRERWTSIRKTISRRLLSALHESAMLTTFNEVYMTPLFHLRKE
KÖEBFLSRYGVKLGFMSFFVKAVLEALKAYPRVNAYIDGEEIVYRHYYDISIAVGIDRGL
VYEVIRDCDKLSNGEIEQKLADLALRAREGLLAIAELEGGGFTITNGGVYGSLLSTPIIN
PPÖVGILGMHKIEKRPVVLDNEIVIADMMYVALSYDHRLIDGKEAVGFLVKVKEGLENPA
SEEBT.

CEMALO378 426195 423445
SucA-Oxoglutarate Dehydrogenase
IVFIERNYFMDSFYGOVYSSOMDMIESHYQFMNHETLDPSWKYFFEGYQLGQAASPSE
ASYKISGNETIAMLQQKSQFLCTIYRYYGYLQSQISTLAPTIDSRFIQEKIAKIDLDEQ
VPSAGLLPKAQVSVRELIEALKKCYGGSLTLETLTCTPELOEFWNILMEKRQVERFABOL
LBSYKDLCKATFFEEFLO IKFTOGKRFSLEGGETLVPMLEHLWHGSALGISNYVLGMAH
RGRENVLTNVLGKPYRYVFMEFEDDPAARGLESVGDVKYHKGYVLKSHQKDRETTFVMLP
NASHLESVDPIVEGVVAALQHQGHAGKEQSSLAILVHGDAAFSCGGVVFETLQLSRYPGY
STEGTLHIVNNYIGFTAVPRESRSTPYCTDIAKHGIFVFRVNSEDVVACIEALEYALQ
VREBFSCDVIIDLCCYRKYGHWESDDPSVTAPLLVDQIKRKKSIRELFRQVLLEGGPADI
STETLASLEFFLOSSLENDESUR DREFOUNKETDEPBEPBKKECHHODEN MAGELLHTODUSLDRET VREEFSCDVIIDLCCYRKYGHNESDDPSVTAPLLYDQIRRKKSIRELFRQYLLBOQFADI
SEETLASIEKEIQESLNREFQVLKGTDPEPFPKECHHCDRLNNGELIHDCDVSLDRET
LEMSSRLCGFPDNFHPHEKIKTLLEKRNKMAEGGVGYDWAMAEELAFASLLIBGYNLRL
SGODSIRGTFSQRHLWWSDTVTGDTYSPLYHLSAEQGSVEMYNSPLSEYAILGFEYGYAQ
QAERTLVLWEAQFGDFANGAQIIFDQYISSGIQKNDLHSDIVLLLPHGYECQGPEHSSSR
IERYLQLAANWNFQVVLPSTPVQYFRILREHAKRDLSDPLVIFTFKLLLRYPQCVSSIEE
FTEPGFRAILEDADPNYDASILVLCSGKIYYDYAEMLPQDRRKDFSCLRIESLYPLALE
DLYSLIDKYSHLKHFYWLQEESKNMGAYDYMFMALQDILPEKLLYIGRPRSSSTASGSAK
LEGGEFLYNFTFIFEIG LSROELVTCMETLFSLR

CPH_0379 CTTC3379 426268 426765
CTTC33 hypothetical protein
KNKKMLCTCSRIODCNPMMSERLKKLESELHDLTOMMOLGLVPKKEISRHQEEIRILEH
KIYEEKERLQLLKENGEIEEYVTPRRSPAKTVYPDGPSMSDIEFVEPTETEIDIDPGTV ELELTDEGREDGAVEVDYSHEDDEDPFSDRNRWRRGGIIDPDANEW

CPn_0380 426671 427876
hemN-Coproporphyrinogen III Oxidase
KSTIPTKTMKTLSAIAIAGDAVVSLIPMLMNGKAPLALYIHIPFCTKKCRYCSFYTIPYK
SESVSLYCNAVIOGEGKKLAPIQETHFIETVFFGGGTPSLVSPLDLKRILKELAPHAREI
TLEANPENLTVSYLRQLQETPINRISVGVQTFDDSILQLLGRTHSSSAAITALQECQNHG
FSNLSIDLIYGLPTQSLEIFLSDLHQALTLPITHISLYNLTIDPHTSYKHRKILVPTIA
GEEILARMSLLAENLLLSQGFQRYELASVAKPDYPAKHNLYWTDRPFGGJVSASOYLH
GERSKNYSHISHYLRAVRKNLPTQETSEILPKKERIKEALALRLRLLEGAJLAEFPSTLI SMLTQDVKLQNLFSVHGQCLALNRQGRLFHDTIAEEIMGYSF

CPn 0381 429836 428037 CT126 similarity
SLPNKFRALMTAPTESRSSPPTLLEETEPLSPNPIPADIQIPRITISPPSLDVSTVASSA SLPHKFRALMTAPTESRSSPPTLLEETEPLSPNPIPADIQIPRITISPSLDVSTVASSA
EDIGYFIAGGPRSSSSASASDVELVCLCGGDEDEPPDSEVRTLYVNGSWOTHQEAVQ
ELLYISEVRGEAVRLLYNDGSGMSPWPISPCRTLPTLDHPLCQALLTVWEGFFSAPENQN
REFLVIFYGDASPYIQQALTQSRHSPRIVVVGISPTVFIQGDFRYMNYRVSGDFPSSLDC
RGTPAENTTILPYSSGLEGVFLPSIRCPSFTWAVRFGEQCLWANGEDVEDRGCLSQDAE
RSQLPHGERDLAVVIDSTDPSSMSRLVEWLNQGSPSSDMEINPYPQRCPDVALSALYAIS
RVJJLAQEWILASVHEGLDLDICYSLILMHTTFAVRYFFLLFTWYPQSREPRTRATRVAQ
OFVYGORVIGGGLPVGTVVRAKLWNPOEDILRAIFISASTTIGGSTYFVECTRWWGRGLRHRVQ
OFVYGORVIGGGLPVGTVVRASYRDRAGFIIGFLQTVHGGLYLDVGIMVLNQIAIQVPRILV
PHIRTTAVYDLHNKSAEENMSSGDVLAVCQTLNFILCAFVLFYNLWFFVKSVLRHSRRRRR

430752 430036 COM_0.9482

4307/52

4303/64

FUTUYLLENTIATRAVETLESVIGELVIRLDGLIVESDEGGRAFLSLWKIPEVHKEPLAI

LUBHARDIYKAWDEYLEFIVKHGENWGLISDAGLECIADFGASLVYRRARADIIPVOARSGE

CSITTLAI MUSGLERGOFFELVY HOEDRWGLISDAGLECIADFGASLVYRRARADIIPVOARSGE

CSITTLAI MUSGLERGOFFELVY HOEDRERKYKSIKKAYTSKEVITSVISVIETVRNYTFE

CLLDFLPSYARLAVADDLRGPSELVLTROVOSWRTT FOLGOVKOSITKVPT I FLEHIPN CPT_0383
CTO47 hypothetical but the CTO47 hypoth KNNVQDP IVAVETLV IRMVNL

13,549 1 -4 hctB-Histone-like Protein 2 ViTCLIRGIKMIGAQKKOSGKKTASRAVRKPAKKVAAKRTVKKATVRKTAVKKPAVRKTA AKKTVAKKTTAKRTVRKTVAKKPAVKKVAAKRVVKKTTAKRAVRKTVAKKPVARK TTVAKGSPKKAAACALACHKNYKHTSSCKRVCSSTATRKHGSKSRVRTAHGWRHQLIKM

414042 432522 CPn_0385 4 #4042 432522

pepA_Leucyl Aminopoptidase A

FLVIKGEFVVLFHAQASGANRVKADAIVLPFWHFKDAKNAASFEAEFEPSYLPALENFQG

KTGEIELLYSSPKAKEKKIVLIGLGKNEELTSDVVFOTYATLTRVLRKAKCSTVNIILPT

ISELRLSAEEFLVGLSSGILSLNYDYPRYNKVDDRILETPLSKVTVIGIVPKMADAIFRKE

AAIFEGVYLTRDLNNRADEITPKKLAEVALNLGKEFPSIDTKVLGKQAIAKEKMGLLLA

VSKGSCVDPHFIVVRYQGRPKSKDHTVLIGKGVTFDSGLDLKPGKSMLTMKEDMAGGAT

VLGILSALAVLELPJWTGIIPATENAIDGASYKMGDVVVGMSGLSVEIGSTDAEGRLI

ADAITYALKYCKPTRIIDFATLTGAMVVSLGEEVAGFFSNNDVLAEDLLEASAETSEPLW

RLPLVKKYDKHJGLADMKNLGSNRAGAITAALFLQRFLEESSVAWAHLDIAGTAYHEK

PEDDDVBVVATGHFATILYYIFNSLSK EEDRYPKYASGFOVRSILYYLENSLSK

434543 434046 CPn_0386 CPT_U386 434543 434046
ssb-SS DNA Binding Protein
KSKGYLMMFGAFAGYLGADPEERMTSKGKRVITLRLGVKTRVGMKDETVWCKCNIWHNRY
DKMLPYLKKØSGVIVAGDISVESYMSKDGSPQSSLVISVDSLKFSPFGRNEGSRSPSLED
NHQQVGYESVSVGFEGEALDAEAIKDKDMYAGYGQEQQYVCEDVPF

CPn 038 435229 434699 CPT_038/
TOTAS Mypothetical protein
NNNNLLGOSIMSRONAEENLINFAKELKLPDVAFDQNNTCILFVDGEFSLHLTYEEHSD
RLYVYXPLLDGLPDMYORKLALYEKLLEGSMLGGGMGGGVGVATKEQLILMHCVLDMKY
AETNILKAFAQLFIETVVKWRTVCADICAGREESVDTMPQMPQGGGMQPPPTGIRA

CPN_0388 435323 437320
glgk-Glycogen Hydrolase (debranching)
stylekvssypsyplelgaskispnryrfalyasoatevilaltdensevievplypdthr
tdalwhielegisogssyafrungpkkhgmgysfkeyladpyaknihsposfgsrkkgod
yafcylkeepfpwoodoplhlpkeemijyehhvrsftossssrvhapgtflgiiekidhl
kklginavellpifefdetahpfrnskfpylcnymgyaplnffspcrryayasdpcapsr
Eefktlyktlhqegievildvvfnhtglogttcslewidtpsyyildaoghfthysogont
Lntnrapttomildilrwweemhvdgfrfblasvfsrgpsgsplofapvleaisfdpll
astkiiaepwoagglyogypflsprwsempcyrdnykaflngonligtfasrisgs
DDIyphgsptnsinyvschdoftlcdtyfynkkheangednbrdtdhysynfotfsekt
EDPGILEVRERQLRNFFLTMVSQGIPMIQSGDEYAHTAEGNNRWALDSNANYFLWDQL
TAKPTLMHFLCDLIAFRKKYKTLFNRGFLSKKEISWVDAMGNPMTWRPGNFLAFKIKSPK
AHVYVAFHVGAODLATLPKASSHEIPVOIVAESOOGFVPONTATFVSLOPHTLIAIS ahvyvafhvgaqdqlatlpkassnflpyqivaesqqgfvpqnvatptvslqphttliais NAKEVT

Pn_0389 438254 437319

CT041 hypothetical protein

TVFNFKRFYQKDSQRQNGNTTCLRPFKKTCKELIEFRRTVKLLKNVLLGLFFSMSISGF
SEVKVSDTFVKQDTVVEPKIRVLLSNESTTALIEAKGPYRIYGDNVLLDTAIQCQRCVVH
ALYBGIRWGEFYPGGQCLKIEFVDDTASLFFNGIQYQGSLYVHRKDNHCIMVSNEVTIED
YLKSVLSIKYLEELDKEALSACIILERTALYEKLLARNPQNFWHVKAEEEGYAGFGVTKQ
FYGVEEAIDWTARLVVDSPQGLIIDAQGLLQSNVDRLAIEGFNARQILEKFYKDVDFVVI

CPn_0390 439171 438134
ruvB-Holliday Junction Helicase
RKSDREGSYMTHOVAVLHODKKFDVSLRFKGLEEFYGGHHLKERLDLFLCAALQRGEVPG
HCLFFGPPGLGKTSLAHLVAYTVGKGLVLASGPQLIKPSDLLGLLTSLQEGDVFFIDEIH
RMGKVAEEYLYSAMEDFKVDITIDSGPGARSVRVDLAPFTLVGATTRSGMLSEPLRARFA FSARLSYYSDODLKEILVRSSHLLGIEADSSALLEIAKRSRGTPRLANHLLRWVRDFAQI REGNCINGDVAEKALAMLLIDDWGLNEIDIKLLTTIIDYYQGGPVGIKTLSVAVGEDIKT LEDVYEPFLILKGFIKKTPRGRMVTQLAYDHLKRHAKNLLSLGEGQ

439701 439510 No robust homolog present in Genebank/EMBL as of 11/7/98 KDQLYKQEKPIPKATILSRNLEVMLDNPKGKRQTLFLGRTSGRSALYSYSRRILVLLNAF

439814 440383 CPn 0392 dcd-dcTP Deaminase
MSIKEDKWIREMALNADMIHPFVNCQVNVNEETGEKLISYGLSSYGYDLRLSREFKVFTN
VYNSVVDPKCFTEDIFISITDDVCIVPPNSFALARSVEYFRIPRNVLTMCIGKSTYARCG I IVNVTPFEPEWEGHVTIE ISNTTPLPAK IYANEGIAQVLFFESSTTCEVSYADRKGKYQ KQQGITVPCV

CPn_0393 440229 440723 CT039 hypothetical protein Ketlehecorkftlakgiprgysisivpparflagtekesiksnkaspylvskvgvrkkn WGFRLLEEVMIKSWWIFSILIGGFV/DRAIDELTEELRLQSKVSSLCODILSAQEKQR QLQLHLQHWQDSAATEAALTORLGLTPKGYKKLCVSPKQQGENKD

440727 441968 CPD_0.394 440727 441978
ETYCHOR DOMAIN PROCESS (HERMOTYSIS HOMOTOG)
KETMIPTMEMPFI ICPTICOSOFICICO (ALECCIPTALICHYKROKSKKOORVATLLEHPH
HELITEIPOIGENIA (NOCATLEGOALEMATTALICHATTILICAETLEKAVALPPHTO
LACOVAPLILAVTK (FREILHMOTYSISTYVVMULGIKKO) (DI TQLOGEKEVLOCKOPPTO
VNOCESPILLAVYLSESDCSVKERMOPPOLICHYD (TYTHENDYLLEIKO) (ICRUPTENTO
LONLOSICTARSILLHIDKPLOGODDLICHEKEPTYMHETTCARMALOMAAEDETIAMIT
DEYGGIEGGIPOEDFETVAGETVDOPOMKILTMINITATICAMITELEREPSETPOINIL
PTNNNITATICAMITEQICT (PTYTMKILTMINITELEVAARRIRIRRYYTIKLYD) CPn_0195 441955 443175
CT257 hypothetical protein
CNCMTNSALFMIGOVIICIVLOGFYSMMEMACVSFNRV / LTKDHKKARYINFLIRR
PYRLFGTVMLGVNILGUVGSESSRNCYRALGITPDVAPFTOIFIVVIFAELLPLTISRKI
PEKLALMGAPILYYSHYIFYPLIQLIGSLTEGLYYLLNIRKEKLNSTLSRDEFOKALETH
HEEQDFNTIATNIFSLSATCADQVCQPLEQVTMLPSSAAVKDPGTIKYTDINFIPVYHK
ARKNVIGIAHEKDFVNKALDEPLINNLHSFWFITAKSKLIR ILKEFRDNRSSVAVVLNS
GEPIGILSLNAIFKILFNTTNIAHLKPKTISVIERTFPCNSRIKDLQKELDIQFPQYPVE
TLAQUOTAGLASPARY/TGVVIINHALGEFYKENLLGGINTVGIKNLLG

443241 444159 CPH_0.396 444.59 44.3241
yhfo-nifs-related protein
ysmiyldnnamtppergllefloktfliegtyanpssvholgkksrolvleashwmokvl
sfogrvlytsgateslnlaiaslpkdshvitsgsehpaileplkhsslsvsylnpedgrc
vltiegieravtpktsaiilgwnsetgakadiaaiahfadergloffvdatawkkeri
vlpsgvtmaafschkfhalsgigalluspgvklheplwsgogogglagotfungiasll
yifkyldlhoerisgeilthrngfekaikaripdvhihcadoprannvsaiafpplegev
loialdiegvacgygsacssgatapfkslvsmgvdeeltlatlrfsfshllloedverav
Glifkyldbilkis

445124 444381 CPT_039/ 445124 444381
PPZC phosphatase family
EHFVDFDYFGLSDIGRVRARNEDFWQVNLMSQVVAIADGVGGRLGGDIASQEAVTSLMEL
IDEQGKLMGYGDDGYKETLKKILLEVNLVYEHGGMEEHLQGMGTTLSFIQFRKDRAWL
FHVGDSRIYRIRGGELRRITEDHSLENQLKNRYGLPKQSDKVYSYRHILTNVLGSRPYVM
PDIRNLPCEKEDLYCLCSDGLTNMVPDIDIRDILNQPATLEERGNALISLANTRGGDDNA TVVLVRIQ

GIIEKVVERLKNS

CPn_0398 445518 445700
No robust homolog present in Genebank/EMBL as of 11/7/98
IEELPMQIENSSILFAEVVMKWFIFSVISAPVVFLPGCTLIPKEKVTKVPSQLWSESLSQ

CPn_0399 445759 446523
CT253 hypothetical protein
YKLMRVLMGKSLNCESIDLKSKNFPRARIFCKISNLRTVTMRKMLVLLASLGLLSPTLSS
CTHLGSSGSYHPKLYTSGSKTKGVIAMLPVFHRPGKSLEPLPMNLQGEFTEEISKRFYAS
EKVFLIKHNASPGTVSGPYAPIANRLPETIIEOFLPAEFIVATELLEQKTGKEAGVDSVT
ASVRVRVFDIRHHKIALIYQEIIECSQPLTTLVNDYHRYGWNSKHFDSTPMGLMHSRLFR
EVVARVEGYVCANYS

CÉTIO400 446527 447306 CTZS4 hypothetical protein SKEMSKFILLLSLGVAALASKNFFIWPAPSGKTPLKLRQVLFGGALLVFSSLVALSVSSQ TAELSTMTGISLAFFLEYLLFLERDITRAILFSGERPVKTSWRALGSATRMMIIIFV TOFIGIMMSKFLTLVLPTQEIHTQEVTQEVQNSLPITGHYISMIINLGVLTPFGEEVFFR GILQTFLKNMTRIAAVLCSSIIFSFIHIEHSLGSWVFVPVLFVFSLSAGFLYEKDRHIL SPIALHGLFNLTSLLFLGIK

CPn=0401 CPT-0401 447884 447495
CT255 hypothetical protein
MTDHAFSKLIGTVRAMVVEGRCPWSLQQSLVSMVEHILGECQEFHEAVLQGKTVQEVGSE AGDVLTLVLILCFLLEREGVLASEDVANEAMEKLRRRAPYIFAEDYKPVSIEEADRLWEL

AGUUTTVLILLERGYLASEDVANEAMEKLRRRAPYIFAEDYKPVSIEEADRLWEL
AGUEEKNEST

CPT_0402

449012

447888
mikY-Adenine Glycosylase
NRKFCMTKIAFSEKAKNFPVEALKKWFEKNKRSLPWRDNPTPYSVWYSEVMLQOTRAEV
VEDYFNQMMERFFTIESLAAAKEEDVIKLWEGLGYYSRARHLLEGARMYMEEFHGKIPDD AESLAO TRGVGPYTVHAILAFAFKRRAAAVGGNVLRVLSRIF LIETSIDLESTRTWSRI AGĀLPHKSPEVIAEALIELGACICKKVPQCHRCPVRQACGAWRENKQFVLPVRHARKKV IFLHRLVAIVLYDGSLVVEKRRFKEMMAGLYEFPYIEVEPEEGLQDIEGFTKKMELSLES PÜEFLGRIKEQRHAFTNHKVHLCPIIFKATSLPQFGELHLLSDIDHLAFSSGHKKIKDAL LIĒKEGDVRSRESIGV

CPT 0403 449009 449710
yce_-predicted pseudouridine synthetase family
yreG_-predicted pseudouridine synthetase family
NPMG_SNDKRAALOYFMENFSWLATQVSRLSSFLRSQLPNHSKQEILASIRQHRRRVNGF
IERFESYKVQOGDRVSLSIIPSTKQOPSILWEDDYSIIYEKPPHLTTEQMAHMTRFFTVH
RLDKGTSGCLLMGKSKQAATELMKLFKQRKIHKQYIAFVFGHPKKKFGTVKSYYAPVYRR
CGAVIFGAAGPSQGEPIKSAYKWDCWVILLSEMSTTDLKNSLPRSSALSSMLTP

CPn_0404 450962 449871
No robust homolog present in Genebank/EMBL as of 11/7/98
ELEALEQKYGKAVLLIALSELGIDTMSLLSGHRLEGFPPIAEVMAACDROSMDFCEILKS
QSMDLWADAASCVUGLLODFFWSTAIASGIAKSSLQETEFECESKYMVLSSWGEGAQVC
SPFNLERICMSFPSLKVFSLKKNGCENMGIQLSASCMNLLMSIFFVATWGGSTPIWITKE
NLMALVALVLSHYQCYFVPATGDPORGNILGNPEVNAILARGMGNRVVLERKRGGESSSS
RYLELAARCFENSLTKTSLLSDANNVQERDKCLLOMSTSLMHTAGLNLQRPPVPTPSGVT
AHPOPQPDPVVTSQPSLLGARERSPVSSRGRFPVVLPLSVISPRSHFGRVERRDLEDEEE

451814 450966 CPI_0405 451814 450966
CT105 hypothetical protein
NIQTSHSRVLLKKFSKEFTIRTYRSLGFTDYLCGCLTNPLGKFPSPQNPQVVTIAPSSTT
PQAVSSAVQGFLQTGCAASSTATTTTASCASALGLISPDQVQALLTNLLNVCQPSVGQPST
SACTSGASSSASMQQGOLLQLILLKTTGSGGSSVSSEQLOQVLSLVSQMTTSQGSGGTQ
ACQAASVLLNLLSATCSAANPLGTAASLAQIIYAAVTSPGAKKTSEFCYNYCGETCGGN
LGCPTCGCPDGQCGCGGGFRFCGVWKNCCGIGEGSQEPAFPL

CPn_040h 451960 452865
tabf-Enoy1-Acy1-Carrier Protein Reductise
CGEMICK (DLTCKVAFVAGLGDDGCYGWOLAKLLAEAGATIIVGTWVPIYKIFSQSWELGK
ENECKKLCNGTLLEIAKIYFMAAGFDGFEDVEDIALMARYKGITGFFIGEVAEQVKKDF
GHIDILVHGLANGPETGKUGLETGRKGYLAALSALCYOFVSLLEHFGSIMBROSTIGHT
YLAHMRAVPGYXXXMGCAKAALECDTKTLAWEAGRAWGIRVWTESAGPLASRAGRAIGFI
ERMYDDYYJEWAPIPEAMNAEQVGAVAAFLAGPLAJAITGETLYVDHGANVWGIGPEMFPK
ES

CIn_0407 453757 45285 HAD supertamily hydrolase/phospitase NYGDAMEKLLVTDID/TTITT DKKVYERLYALHQA/WKLFFUTGP/YKYAAPLFSD FDAPYLLCQNCASVWGTT. SKSJØFSDLLGILQDCMEGATALFSVESGAPYGDHY YRFSPTPIAQDLHEYVDPRYFPNÁKERB/LFETRSLKDDYAFPSFAAKKYGLRDEVIRI GKELEROBALTSVATMTLMRWFFDRYÅLLFLTDKSVSKGALDRVVNILYDCKKPFVMA SGDDANDLDLIERGDFKIVMSSAPEENHVHADFLAPPADKNGILSAWEAGVRYYDDLMSL

454020 454581 CT102 hypothetical protein ORDINGRAMED SOMKLEMERMICOLLILL TT CORP IN LITE AND MAIN BAND OF THE COLLECTION OF THE COLUMN TO A STREET COLUMN TO A STR ATSOFLSKATUKTURKQQTNQD **EFPPDTDINHLLQENIKQSSSS**

454645 455127 CPn_0409 CT160 hypothetical protein
mTTWTLNQNNLTKFLKSSD#EPFLERESGLTYINIQANGNELPLFFVIRSEGEILQLICY
LPYQLHESHKASTARLLHYLNRDIDIPGFGMDEEQGLIFYRLVLPCLNGEIHDTLLRIYI DTIKLVCDSFSHAIGLIS GOMNLDELRROALQEQQEKRNE

455087 CPn 0410 455833 dnaQ-DNA Pol III/ Epsilon Chain
DVRLFKSNKKNVMSSØTMDVLIFYDTETTGTQIERDRIIEIAAYNSVTDESFLTYVNPEI PIPDEASKIHGITTÍÁVLSAPKFPEAYEGFRKFCGEDSILVANNNOGFDEPLÍCKEREN PIPDEASKIHGITTÍÁVLSAPKFPEAYEGFRKFCGEDSILVANNNOGFDEPLÍCKEREN SLEPLTNRTIDSILVAAOKYRPDLPKHNLOYLROVYGFAENQAHRALDDVVILHKVFTSLI GOLPPQOVLDLLOSYHPKVFKMPFGKYKGQPLVDIPKSYFEWLENQGALDKPENKDIKA AIALLHQPT

CPn_0411 455794 456609
CT262 hypothetical protein
RHOSRYSSITSTDMILITAAFSPCPNDIFLFRSFLKDPOFRPLLNOVTIADIETLNTLALO
RRLSLMKNSAALFFLVSDYYNLMDVGNTLGYNSGPIVLSLDPECSLDTLATPGEMTTAHA
LCKLYYPKAKLIPMPVDKILSAILOGKVOGGALHHERFSYDLQLTLRADFGELWRRKTI
FPLPLGCIAIAKYVPMATVDALTAALRKSLICSLKDPITAGAKAVEYSKNKNVTVIHRFI
GTYINKDFFQLSKTGKKALHHLWKANECCOYT

CPn 04/12 456515 457246 CFILOUIZ

CT263 hypothetical protein
EPISTKKFFNYLKLGKKLYICSGRFMNAVNTFKKILCIVADYREISPLIEQLDFTOINEH
LYSYRCTDYHLDLYIVHVWGSTAVLNALOSYCOAYTDYDLWINPGFVGACSPEIPLGCCY
TIGKIANLTTDTPEVLSEDPPYIFDALPDSLPKSSLVTSFVLYHYGFHKTFKLLDMDDYA IASOAAEHHIPCSFLKITSDYTVPGDCPFSRLEEVSQKLTOTLVELLPELMERAIPPKLL LPCP

Pn_0413 459209 457227

msba-transport atp Binding Protein

VFMKLLLKAVLRHKNHLVILGCSLLAILGLTFSSQMEIFSLGMIAKTGPDAFLLFGRKES
GKLVKVSELSGKDILENWOJAISKDSETLTVSDATTYIAEHGKSTASLTSKLSKFVRNYID

VSRFRGLAIFLICVAIFKAVTLFFGRFLGGVVAIRVSRDLRGDYFKALQUEPMTFFHDID

IGNLSNRVMTDSASIALAVNSLMINYIQAPITFILTLGVCLSISWKFSILICVAFPIFIL

PIVVIARKIKALAKRIQKSODSFSSVLVDFLAGVMTVXVFRTEKFAFTKYCEHNNKISAL

EEKSAAYGLLFRPLLHTIASLFFAFVVVIGIYKFAIPPEELIVFCGLLVLIYDPIKKFGD

ENTSIMRGCAAAERFYEVLNHPDLHSQKEREIEFLGLSNTITFENVSFGYQEDKHILKNL

SFTLHKGEALGIVGPTGSGKTTLVKLLDFLKEVSGGKLIDDSLPITEYNKGSLRNHACV

LQNFFLFYDTVWNNLTCGKDMEEEAVLEALRRAYADEFILKLPKGYHSVLEESGKNLGC

QQQRLAIARALLKNASILILDEATSALDAISENYIKNIIGELKGQCTQIIIAHKLTTLEH

VDRVLYIENGQKIAEGTKEELLQTCPEFLKMWELSGTKEYNRVFVPDHKLVANPTDMAIT

T

460203 459172 CPH_0414 460203 459172
accA-AcCoA Carboxylase/Transferase Alpha
LCLRIVCIKMILFIRGEHILMELLPHEKQVVEYEKAIAEFKEKNKKNSLLSSSEIQKLEK
RLDKLKEKIYSDLTPWERVQICRHPSRPRTVMYIEGMCEEFVELCGDRTFRDDPAVVGGF
VKIQGQRFVLIGQEKGCDTASRLHRNFGMLCPEGFRKALRLCKLAEKFGLPVVFLVDTPG
AYPGLTAEERGOGWAIAKNLFELSRLATPVIIVVIGEGCSGALGMAVGDSVAMLEHSYY
SVISPEGCASILWKDPKKNSEAASMLKMHGENLKQFGIIDTVIKEPIGGAHHDPALVYSN
VREFIIQEWLRLKDLAIEELLEKRYEKFRSIGLYETTSESGPEA

CPn_0415 461522 460221
CT266 hypothetical protein
SQTGFLPGLTLIFVIIIWCNAFLIKLCVIMGLQSRLOHCIEVSQNSNFDSQVKQFIYAC
QDKTLRGSVLKIFRYHPLLKIHDIARAVYLLMALEEGEDLGLSFLNVQQYPSGAVELFSC
GGFWKGLPYPAEHAEFGLLLLQIAEFYEESGAYVSKMSHFQQALFDHQGSVFPSLMSQE
NSRLLKEKTTLSOSFLFOLGMQIHFEYSLEDPALGFWMQPTRSSSAFVAASGCQSSIGAY
SSGDVGVIAVQPGSGDISDCYYFGCCGIAKEFVCQKSHQTTEISFLTSTKDKPHPRNTGFS
YLROSYVHLPIRCKITISDKQYRVHAALAEATSAMTFSIFCKGKNCQVVDGPRLRSCSLD
SYKGPGNDIMILGENDAINIVSASPYMEIFALQGKEKFWNADFLINIPYKEEGVMLIFEK
KVTSEKGRFFTKMN KVTSEKGRFFTKMN

CPn_0416 461871 461557 himD/ihfA-Integration Host Factor Alpha EALSNMATMTKKKLISTISQDHKIHPHHVRTVIQNFLDKMTDALVKGDRLEFRDFGVLQV VERKPKVGRNPKNAAVPIHIPARRAVKFTFGKRMKRLIETPNKHS

463047 462244 CPI_041/ 46304/ 463244
amiA-N-Acetylmuramoyl Alanine Anidase
REKGMKLTKYLNTKOLRSMISRLFVRYSLFMSKOLSFFALCVLGSHPIFAOTPNPPORVR
RSEVIFIDPGHGGKDQCTASKELHYEEKSLTLSLALTVGSYLKRMGYKPQLTRSSDVYVD
LGKRVALSNRCQGDVFISIHCNHSSNAAAFGTEVYFYNGYGSPTRNRMSEVLGKNILAA
MEKNGILKSRGLKTANFVVIRDTSMPAVLVETGFLSNSPERAALQDARYRMHVAKGIAEG
VHNFLCGPSFQKPKQNIAKIRKPQIQAN

466897 454875 CPn 0419

CPH_0419 466897 464876
pbpl-transstyrolase/transpeptidase
QLFPNTNIWNIPQKKVSVFYPMSYRKRSTLIVLGVFALYALLVLRYYKIQICEGDHWAAE
ALGOHEFCVRDPFRRGTFFANTTVRKGOKDLQQPFAVDITKFHLCADPLAIPECHRDEII
OGILOFIEGGYYDDLGLKLDKKSRYCKLYPLDVSYHDRLSLWMKGYATKHRLPTNALFF
ITDYQRSYPFGKLLGQVLHTLREIKDEKTGKAFPTGGMEAYFNHILEGDVGERKLLRSPL
NRLDTNRVIKLPKGTGDIYLTINPVIOTIAEEELERGVLEAKAQGGRLILMNSGTGEILA
ALGOYPEPDFTCHEUFERBERGHERFFERSENDENDEREGGIHKENGTWAALGANEEAGLKG
KFIERFEED ITSTETLEFERBEGGHERG GEBEGGERGTWAALGERGTWAALGANEEAGLKG
VANYQOKLLALGFGRKTGIELPSEASGLVPSPHRFHINGSLEWSLSTPYSLAMGYNILAT
GLOWMAAVAILANGGVAUPDPLIVKKUVASGEFGYHLPTKEKFRIEFSFETIREGVARNIET GIOMVOAYA ILANGGYAVRPTLUKKIVSASGEEYHLPTKEKTRLFSEEITREVVRAMRFT TLRGSGFRASPKHHSSAGKTGTTEKMIHGKYDKRRHIASFIGFTPVESSEGNFPPLVML VSIDPEYGLRADOTKNYMGGRCAAPIFSRVADRTLLYLGILPDKKLRNCDEEAAALKRL YEEWNRSPKQGGTR

CPn_0420 467120 466824
CT271 hypothetical protein
KSFPMNKSRFLRLCCCLCFCGSLFYFYINKQNSLTKLRLEIPCLSVRLRQLEQQNISLRF
LIDKIERPDHLMEIAALPEYQYLEYPSEESISLLSYELP

CPn_0421 468007 467108
yabC-PBP2B Family methyltransferase
EILMSERAHIPVLVEECLALFAQRPPQTFRDVTLGAGGHAYAFLEAYPSLTCYDGSDRDL
QALAIAEKRLETFODRVSFSHASFEDLANQPTPRLYDGVLADLGVSSMQLDTLSRGFSFQ
GEKEELDMRMDQTOELSASDVLNSLKEEELGRIFREYGEEPQWKSAAKAVVHFRKHKKIL
SIQDVKEALLGVFPHYRFHRKIHPLTLIFQALRVYVNGEDRQLKSLLTSAISWLAPQGRL
VIISFCSSEDRPVKWFFKEAEASGLGKVITKKVIQPTYQEVRRNPRSRSAKLRCFEKASQ

468233 468784 CT273 hypothetical protein GLAWVEIFNYSTSIYEQHASUNRIVSDFRKEIQMEGISIRDVAKHAQILDMNPKPSALTS LLOTNOKSHWACFSPPNNFYKORFSPYLAPSLGSPDQQDEDIEKISSFLKVLTRGKFSY RSQITPFLSYKDKEEEEDEDPEEDDDDPRVQQKVLLKALDLGVKSTNVMIDYVISRIFQ

468788 469216 CPI_0423
400720
CT274 hypothetical protein
CMLDNEWKAILGWGDDELEELRISGYSFLRQGHYSKAILFFEALVILDPLSIYDHQTLGG
LYLQIGENSQALAVLDQALRWQGDHLPTLLNKTKALFCLGRIEEATAIATYLSSCPIPAI ANDAFALLMSYSKATKKNAALVR

CPRE0424 469528 470961

char-Replication Initiation Factor

SREEIFSPSLMGWVDCIWESFINKESGMLTCNECTTWEQFLNYVKTRCSKTAFENWISP

IQVEETOGERIRLEVPINIFVONYLLDINYKRDLCSFVPLDVHGEPALEFVVAEHKKPSAPV

ASQKESNEDISEVFEETKDFELKLNLSYRFDMFIEGPSNOFVKSAAVGIAGKFGRSYNPL

FIEGGVGLSKTHLLHAVGHYVEHHKNLRIHCITTEAFINDLVYHLKSKSVDMKNFYRS

LDLLLVDDIQFLQNRGNFEEEFCNTFETLINLSKQIVITSSKPSQLKLSERIIARMEMG

LVÄNVGIPDLETRVAILOHKAEQKGLLIPNEMAFYIADHIYGNVRQLEGAINKLTAYCRL fgksliettvretlkelfrspykoxisvetilksvatvfovklndlkonsrskolvlaro Iamalaktlitoslvaigaafgkthstvlyacktiehklondetlkrovnlcknhivg

CPRE0425 470965 471564
CT276 hypothetical proteins
FRGCPMFRRTGKGFFEDVOTLYEETTSSPSYSPYSRSERPETPPSLFDNPKASEARPLN
HNLTEESSLPQWSSTPRTESLLPLEEPETTLGGVTFKGELAFERLLRIDGTFEGILVSK
GKEGFKGVVKADIOLOGATIEGVVEGNITVSGKVELRGGAIIKGDIQANTLCVDEGVR
ILGYLAIAGITDHSERERDL

CPn_0426 CT277 similarity 472111 471536 WYEFSLIFPKLCYGCQAPGAYFCSNCLEKLLVEDREGRCLHCFRYLGSSETRLCSQCSPS SQEGAFSLYLPSGTALSVYARACEGKRPALQFFSKSIAFELASLDETPSCIAYITSTISR KIYYEVAKLEKLLRIPLWPWLPKKRQIEKLPKGEGICFLSAYPLSQKWMQTIVGGSASPL VSTGEFLSONDQ

CPREGA27 472153 473715

nqr2-NADH (Ubiquinone) Dehydrogenase
AVCYYFERVEASTFLSITMLKKFINSLMKLCOODKYQRFTPIVDAIDTFCYEFIETPSKP
PFIRDSVDVKRWMMLVVIALFPATFVAIWNSGLQSIVYSSGNPVLMEQFLHLSGFGSYLS
FVYKEIHIVPILWEGLKIFIPLLTISYVVGGTCEVLFAVVRGHKIAEGLLVGGILYPILL
PPTIPYWMAALGIAFGIVVSKELFGGTGMNILNPALSGRAFLFFTFPAKMSGDVWVGSNP
GVIKDSLMKMNSSTGKVLIDGFSOSTCLOTLMSTPPSVKRLHVDAIAANMEHIPHVPTQD
VIHSQFSLWTETHPGWVLDNLTLTQLQTFVTAPVAEGGIGLLDFOFDSAFAITDVIYGG
KFSAGNLFWGNIIGSLGETSTFACLLGAIFLIVTGIASWRTMAAFGIGAPLTGWLFKFIS
VLIVGONGAWAPAAFFIPAYRQLFLGGLAFGLVFMATDDVSSPTMKLGKWIYGFFIGFMT
IVIRLIMPAYPEGGWLALLLGNVFPALDYFFAVRYRKGV CPn 427 IVIRLINPAYPEGVMLAILLGNVFAPLIDYFAVRKYRKRGV

474681 473719 CPI_0428

A7379

A74881

ingr3-NADH (Ubiquinone) öxidoreductase, GammaNMSKGSSKHTVRINOTWYIVSFILGLSLFAGVLLSTIYYVLSPIØEQAATFDRNKQMLLA
AHILDFKGRFQIOEKKEWYPATFDKKTQLLEVATKKVSEVSYPILELYAERFVRPLLTDA
OGKVFSFERNLNPIEFFEKYÖESPEÇOSPLPFYVILENTSPTENMSGADVAKDLSTVQ
ALIFPISGFGLWGPIHGYLGVKNDGDTVLGTAWYQOGETPGLGANITMPEWQEGFYGKKI
FLQDSSGTTNFATTDLGLEVVKGSVRTTLGDSPKALSAIDG SGATLTCNGVTEAYVOSL
AVYROLI INFENITHEKKTGF ACYROLLINFSNLTHEKKTGE

474666 CPT_0429 474666 475319
mjr4-NADH (Ubiquinone) Reductase 4
KENNEMMOKKOYKSYFFDPLWSNNQILIAILGICSALAVTTTVQTAITMGIAVSIVTGCS
GFFVOLLRKFTPDSVRMITQLIIISLFVIVIOQELKAFFFDISKTLSVFVGLIITNGIAVSIVTGCS
GROEDLARHVTPIPAELDFRAGGISVGWVLLVIVGVIJELFGGGTLMGFRIIFOFVYAGET
HPD/YQHLDIMVLAPSAFFLLGIMIWLANIRDSKKFKR

CDTPPHILE WITTER TEMAPHOLIAMAPMOLITY OF SECURITY SERVICES OF SERVI 475323 426.093

SKHOPSISKARTORRSL

476489 476151 n Genebank/EMBL as of 11/7/98 No robust homolog present KIMTTLEKYVPRSRONPOTLTFLKRYSSVLLHSENGLSYRIFAKVLAILLTSLAVAFAVT LESCEGSQLRLCALYIGIALAICVLLTIVVYCIASKIATACKKPPSISRIEIV

3/ 4/6514 An Top General Romming on Collins of Andrew Collins Collins on Co No Friedst trom 1 och Killerbyndpadevarby JUFUTUSAPLOLJ IWIAAGUT

GCSH-Glycine Cleavage System H Protein RTFRILYGTLYRTGSRKVMYSDYHVWILPVHERVVRLGLTEKMQKNLGAILHVDLPSVG SLCKEGEVLVILESSKSA/EVLSPVSGEVIDINLDLVDNPQKINEAPEGEGWLAVVRLDQ DWDPSNLSLMDEE

CPn_0434 479471 477276

CT283 hypothetical protein

RMWRIYOODLFCRLYCRPAMFFSLLSFTLRFYCLGRGWTLLSFFYKHQKKFIGIVIAVV

CVSGIGVOWGRFSRKGSAESTSRRTVFTTASGKRYVEKDFMAMKKFFAHEAYPFTCNPRA
WNFINEGLLTDYFTTRYGEKLFLKVYHPGEKIFSKEKAYQPYRREDAPFISSEEVWKSS

APOLLEILKVPOGIENPISKEGFLARAKLFLEERRFPHYVLRQMLEYRROMFALPPDEAL

SRGKDLRLFGYGTIQDWFGDAYLSAAVELLIRFIDEQKKVLPRPSKQEARDDYDKAKHA
YTKISKNKEFSLGFEFVNSYFQFLEISESEFFNMYRDILLCKRALLLLQGGVSFDFOPL

TTFFVQGKDS/OVEFFRLPKEYSFKTKQELKAFEVYLKLVSLPKSDSLDVPNEILPIATI

KAKEPRLVGRFSIDYKRVALQDLAATVFMVEVLHWQONSEHFQEILQGFPDVETCQSYK

DFOHLKPALFOKISLFTRKEILRARPERILOSLQOVPKQSOEVLLSAGKNSALPGISDGQ

QLAKVLLEMEVLDLYSODAETYYTIIVNSSFEKEEVLPYREVLKRDLASQLITSHGHLVD

MERLESAIRTRYPGEEGASLWORRLWKVVENHRLGRHLEGSFSWSLDRSLKTFSRGDKEL

POEFDRIFSMKVGDYSSVFMSPNEGPCYYQCLSHLLYDRPASVDKLFLAKSQLDEELLGS

YMERFIPGGWR 479471 477276

CPn_0435 480908 479475
Phospholipase D superfamily (uncleavable leader peptide)
GVMMSRLRFRLAALGIFFILLVPNSVSAKTIVASDKEKVGVLVYDNSVEAFQQILDCIDH GVMJSRLAFRILAALGIFFILLVPNSVSAKTIVASDKEKVGVLVYDNSVEAFQOILDCIDH
ANF/VELCPCMTGGRTLKEMVDHLEARMDLVPELCSYIIIQPTFTDAEDQKLLKALKERH
PNRFFYVFTGCPPSTSILAPNVIEMHIKLSIIDGKYCILGGTNFEEFMCTPGDEVPEKVD
NPRLFVSGVRRPLAFRDQDIMLRSTAFGLQLREEYHKQFAMWDYYAHHMWFIDNPEQFAG
ACPPLTLEQAEETYFPGFBKHEDLVLVDSSKIRIVLGGPHDKQPNPVTQEYLKLIQGARS
SVKLAHWYFIPKDELLNALVDVSNHNGVHLSLITNGCHELSPAITGPYAMGNRINYFAL
YGKRYPLWKKWFCEKLKPYERVSIYEFAIWETQLHKKCMIIDDEIFVIGSYNFGKKSDAF
DYESIVVIESPEVAAKANKVFNKDIGLSIPVSHGDIFSWYFHSVHHTLGHLQLTYMPA

CPn_0436 481633 480902 lplA-Lipoate Protein Ligase-Like Protein FYVCYMKVRIVDSGKSSAASHMAKDRDLLESLQDGELILHLYEWENPCSLTYGHFMRPEK FLLSNYADLGLDAAVRPTGGGFVFHKGDYAFSVLMSATHPSYSSSVLENYHTVNSFVAKV LEKVFRIGGMLAPEDENSSSRDSGNFCMAKTSKYDVLFGDKKIGGAAQRKVQQGFLHQGS LFLSGSSSEFYQRFLKPEVLEEIIEQIQIHAFFPLGLEAADEVLQEARQQVKEAFIKLFC

CPT_0437 481810 484350

ClpC-ClpC Protease

YMMEKFYNRAKOYIKLAKKEAQRLNHNYLGTEHILLGLLKLGGGVAVNVLRNLGIDFDT
ARQEVERLIGYGPEIOVYGDPALTGRVKKSFESANEEASLLEHNYVOTEHLLLGILHQSD
SVALOVLENLHIDPREVRKEILRELETFNLQLPPSSSSSSSSRSNPSSSKSPLGHSLGS
DKNEKLSALKAYGYDLTEMVRESKLDPVIGRSSEVERLILILCRRRKNNPVLIGEAGVGK
TAIVEGLAQKIILMEVPDALRKKRLITLDLALMIAGTKYRGOFEERIKAVMDEVRKHGNI
LLFIDELHTIVGAGABGGAIDASNILKPALARGEIQCIGATTIDEVRKHEKDAALERRF
QKIVVHPPSVDETIEILRGLKKKYEHHNVFITEEALKAAATLSDQVYHGRFLPDKAIDL
LDEAGARVRVNTMGOPTDLMKLEAEIENTKLAKEQAIGTOEVEKAAGLRDEEKKLRERLQ
SMKQEWENHKEEHOVPVDEEAVAQVVSLOTGIPSARLTEAESSKLLKLEDTLRRKVIGO
DAVTSICRAIRRSTGIKDPNRPTGSFLFLGPTGYGKSLLAQQIAIEMFGGEDALIQVDM
SEYMEKFAATKMMGSPRGVVGHEEGGHLTEQVRRRPYCVVLFDEIEKAHPDIMDLMIQIL
EQGRLTDSFGRKVDFRHAIIITMTSNLGADLIRKSGEIGFGLKSHMDYKVIQEKIEHAMKK
HLKPEFINRLDESVIFRPLEKESLSEITHLEINKLDSRLKNYQMALNIPDSVISFLVTKG
HSSPEMGRAPLRRVIEGVLEDPLAELLLKESCRQEARKLRATLVENRVAFEREEEEGEALK 481810 484350 HSPEMGARPLRRVIEQYLEDPLAELLLKESCRQEARKLRATLVENRVAFEREEEEQEAAL

CPn_0438 485455 484334
ycbF-PF-loop superfamily ATFase
NLTLEMPROVREIMQOTVIVAMSGGVDSSVVAYLFKKFTNYKVIGLFMKNWEEDSEGGLC
SSTKDYEDVERVCLQLDPYYTVSFAKEYRERVFARFLKEYSLGYYFNNPDILCNREIKFD
LLOKKVZELGGDYLATGHYCRLMTELQETQLLRGCDPQKDQSYFLSGTPKSALHNVLFPL
GEMNKTE/RAIAAQAALPTAEKKDSTGICFIGKRPFKEFLEKFLPNKTGNVIDMOTKEIV
GQHQGAH'YYTIGQRGLDLGGSEKPCYVVGKNIEENSIYIVRGEDHPQLVLRELTARELN
WETTDPK/GCHGSAKVMPSDPAE/ATLIVESCDER/MIDESONIAN MTWOOTLEVK WFTPPKSGCHCSAKVRYRSPDEACTIDYSSGDEVKVRFSQPVKAVTPGQT1AFYQGDTCL GSGVIDVPMIPSEG

CPn_0439 485523 486077
No robust homolog present in Genebank/EMBL as of 11/7/98
IISSNNEY/LFVSSTLNGVFPSSLPEESADLFITNKEIVALGEKGNVFLTHSIPMKIAAIT
ILVIVALAGIAIICLGCYSQSILLIAVGIVLTILTLLCLQALVGFIKFIRQLPQQLHTTV OF IREK IRPESSLOLVTNAORKTTODTLKLYEELCDLSQKEFKLOSTLYOKRFELSHKNE

CPn 0449 486081 486740 10 TODIOT HOMOTO: 486081 - 486740

NO TODIOT HOMOTO: PRESENT IN GENEDARK/EMBL AS OF 11/7/98

LATIRGRAMATSVAPSTVPRIOFLOHATEVLNLPHAY (TOPHPIPAAPWETFRSKLSTKH
TILGFALTLLLTIXXT GAGYAGYTMINNI (CHIGUSI IVILTI ILALLLA FILKKKOTCTKL
IDELOGA IOS IGROFVORYGLMETTI KRIVILLELELTTONDEKTR FLINE FAKKEGIONLEL
KITEGGIFLAGKOPKREGGEGFMRG FRHLGKNPYTLFFG.

486875 487838 CPU_0441 ASSEM ADDRESS
CT007 ASPETED ADDRESS
EMBELS/PREFETALE DECEMBER
WATEGOTOFYNYULGE/ATTLOCKIM/WGTT (COLVDEKTIEMOWG)LYGAVULGCKYQAT
EKLSATESY/INETGLHOEKAWEL/YG/CYKATK/UTLICT (FYF/NF6TD/YG/TVC/NLGLAY
EKLSATESY/INETGLHOEKAWEL/YG/CYKATK/UTLICT (FYF/NF6TD/YG/TVC/NLGLAY
EKLSATESY/INETGLHOEKAWEL/YG/CYKATK/UTLICT (FYF/NF6TD/YG/TVC/NLGLAY GEYOWS LOND ESTADDH

489528

CTOUG hypothetical protein NILEGYTHMGFKNICKGOSGLYLHVIFPERILARKLKNCAKSYPRTALTIEVLVSSVLGAL KVILIFCASTYAALTLEDHRALFHAIRTKHSCOHLASYAMAMLLHILTIAVIIGLVFSLVFI PPPVVFISLULLMGVTTSVTLFG/HKNLFPPYEPPSRPHTPPPFADE/VPLISESYFD

4397. 489973

CTOUS Mypothetical protein
USMSOPPINPLOOPQVPAAASPSOPSVVKRLKTSSTGLFKRFITIPDKYPKMRYVYDT
GIIALAAIAILSILLTASGNSLMLYALAPALALGALGVTLLISDILDSPKAKKIGEAITA
IVVPIIVLAIAAGLIAGAFVASSGTMLVFANPMFVMGLITVGLYPMSLNKLTLDYFREH LLRMEKTOCTAPFILVTPSADDAKKIAVEKKDLSASARMEHASORQDARHRRIGRE
AQGSFFYSSRNPEHRRSFGSLSRFKTKPSDAASTRPASISPFFKDDFQPYHFKDLRSSSF
GSGASSAFTFIMPASSRSPNFSTGTVLHEPVYPKKOGKEPSIPRVSSSRRSPRDRQDKQ
QQQQQQQQQQQQQQQQCEQKQQSKKKSGKSQSLKTPPPDGKSTANLSPSNPFSDGYDEREKRKHRKNK

CPn_0444 490266 494507
pmp_6-Polymorphic Outer Membrane Protein
KAFPQRHMKYSLPWLLTSSALVFSLHPLMAANTDLSSSDNYENGSSGSAAFTAKETSDAS
GTTYTLTSDVSITNVSAITPADKSCFTNTGGALSFVGADHSLVLQTIALTHDGAAINNTN
TPALSFSGFSSLLIDSAPATGTSGGKGAICVTNTEGGTATFTDNASVTLQKNTSEKDGAAV
SAYSIDLAKTTTAALLDQNTSTKYGGALCSTANTTVQGNSGTVTFSSNTATDKGGGIYSK
EKDSTLDANTGVVTFKSNTAKTGGAWSSDDNLALTGNTQVLFQENKTTGSAAQANNPEGC EKDSTLDANTGVVTFKSNTAKTGGAWSSDDNLALTGNTQVLFQENKTTGSAAQANNPECC
GGAICCYLATATDKTGLISQNQEMSFTSNTTANTGGAIYATKCTLDGNTTLTFDQNTAT
AGCGGAIYTETEDFSLKGSTGTVTFSTNTAKTGGALYSKGNSSLTGNTNLLFSGNKATGF
SNSSANQEGCGGAILAFIDSGSVSDKTGLSIANNQEVSLTSNAATVSGGAIYATKCTLTG
NGSLTFPGNTACTSGGAIYTETEDFTLTGSTGTVTFSTNTAKTGGALYSKGNNSLSGNTN
LLFSGNKATGPSNSSANQEGCGGAILSFLESSASVSTKKGLWIEDNENVSLSGNTATVSGG
AIYATKCALHGNTTLTFDGNTAETAGGAIYTETEDFTLTGSTGTVTFSTNTAKTAGALHT
KGNTSFTKNKKALVFSGNSATATATTTTDQECCGGAILCNISESDIATKSLITENESLSF
INNTAKRSGGGIYAPKCVISGSESINFDGNTAETSGGAIYSKNLSITANGFVSFTNNSGG
KGGAIYIADSGELSLEAIDGDITFSGNRATEGTSTPNSIHLGAGAKITKLAAAFGHTIYF
VDPITMEAPASGGTIELLUNFVVAK UVPPDQNKAPIASVGPVAVARAPKRTTTIYFS KGGATYIADSGELSLEAIDGDTFSGNRATEGTSTPNSIHLGAGAKITKLAAAGHTIYF
YDPITMEAPASGGTIELEUVINPVVKAIVPPOPPNKOPIASVPVVPVAPANPHNTGTIVFS
GKLPSQDASIPANTTTILNQKINLAGGNVVLKEGATLQVYSFTQQPDSTVFMDAGTTLET
TTTNNTDGSIDLKNLSVNLDDFNIPSSMAAPDYGYQGSWTLVPKVGAGKVTLVABW
ALGYTPKPELRATLVPNSLWNAYVNIHSIQOEIATAMSDAPSHPGIWIGGIGNAFHQDKQ
KENAGFRLISKGYIVGSSMTTPQEYTFAVAFSGLFGKSKDYVVSDIKSQVYAGSLCAQSS
YVLELHSSLRRHVLSKVLPELPGETPLVLHGQVSYGRNHHNMTTKLANNTQKSDWDSHS
FAVEVGSSLPVDLHYRYLTSYPYVKLQVVSVNQKGFQEVAADPRIFDASHLVNVSIPMG
LTHEHESAKPPSALLTLGYAVDAYRDHPHCLTSLTNGTSWSTFATNLSRQAFFAEASGH
LKLEHGLDCFASGSCELRSSSRSYNANCGTRYSF

LKIEHGLDCFASGSCELRSSSRSYNANCGTRYSF

THE CRIMEN AND A 197579

DDF7 - POLYMORPHIC OUTER MEMBRANE PROTEIN

PREFASKKCLOMKSSVSWLFFSSIPLFSSLSIVAAEVTLDSSNNSYDGSNGTTFTVFSTT

DAARGTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGT

VASSABADKNLFNDFSKLSISCEPSLLLSPTQCCALKSVGNLSSLTGNSQIIFTQNFSSD

NGVINTKNFLLSGTSQFASFSRNQAFTGKQGGVVYATGTITIENSFGIVSFSQNLAKGS

GGALYSTDNCSITDNFQVIFDGNSAWEAAQAQGGAICCTTTDKTVTLTGNKNLSFTNNTA
LTYSGAISGLKVSISAGGPTLFQSNISGSSAGQGGGAINIASAGELALSATSGDITFNN

NOVENGSTSTRNAINIIDTAKVTSIRAATGQSIYFYDPITINGTAASTOTHININLADANS

EIFYGGAIVFSGEKLSPTEKAIAANVTSIRAATGQSIYFYDPITINGTAASTOTHININLADANS

EIFYGGAIVFSGEKLSPTEKAIAANVTSIRAATGQSIYFYDPITINGTAASTOTHININLADANS

EIFYGGAIVFSGEKLSPTEKAIAANVTSIRAATGASIYFYDPITINGTAASTOTHININLADANS

EIFYGSTUSTASAKEANLSLNGLAVNLSSLDGTNKAALKTEAADKNISLSGTIALIDTEGS

FENNIKASSTYPLLELTTAGANGTITLGALSTLTLQEPETHYGYGGMVOLSWANATSS
KIGSINWTRTGY ISSPERKSNLPLNSLWGNFIDIRSINQLIETKSSGEFFERELWISGIA

NFEXREDSWFTRHGFRHISGGYALGITATTPAEDQLTFAFCQLFARDRNHITGKNHGDTYG

ASLYFHHTEGLFDIANFLWGKATRAPWVLSEISQIIPLSFDAKFSYLHTDNHMKTYYTON

SI KESSWRNDAFCADLGASLPFVISVPYLLKEVEFFVKVOYIYAHQQDFYERHAEGRAFF

KSELINVELPIGVYTERPGSKSEKGTYDLTLWYILDAYRRNPKCQTSLIASDANWAYGTV

LARGSFSVRAANHFQVNPHMEIFGQFAFEVRSSSRNYNNINLGSKFCF

CPALB446 497602 500415

pmp_B-Polymorphic Outer Membrane Protein
LIERHLSMK1EHKLLISSTLVTPILLSIATYGADASLSPTDSFDGAGGSTFTPKFTAD
ANGTNYVLSGNYYINDAGKGTALTGCFTETTGDLFTTGKGYSFSFNTVDAGSNAGAAS
TTADKALTFTGFSNLSFIAAPGTTVASGKSTLSSAGALNLTDNGTILFSONVSNEANNNG
GAITTKTLSISGNTSSISTFTSNSAKKLGGAIYSSAAASISGNTGQLVFMNNKGETGGGAL
GFEASSSITONSSLFFSGNTATDAAGKGGAIYSEAAASISGNTGQLVFMNNKGLFFAGNSTCGGAICAHGLDLSAAGPTLFSNNRCCNTAAGKGGAIAIADSGSLSLSANGGDITFLGNTLT
STSAPTSTRNAIYLGSSAKITNLRAAQCQSIYFYDPIASNTTGASDVLTINOPDSNSPLD
YSGTIVFSGEKLSADEAKAADNFTSILKQPLALASGTLALKGNVELDVNGFTTTEGSTL
MOPOTKLKADTEAISLTKLVVDLSALEGNKSVSIETAGANKTITLTSPLVFQDSSGNFYE
SHINOAFTQPLVVFTAATAASDIYJDALLTSPVQTPEPHYGYQGHWEATWADTSTAKSG
TMTWYTTGYNPNPERRASVVPDSLWASFTDIRTLOOIMTSOANSIYQORGLMASGTANFF
KKDKSGTNOAFRHKSYGYIVGGSAEDFSENIFSVAFCOLFGKDKDLFIVETTSHYNLASL
YLQHBAFLGGLPMPSFGSITDMLKDIPLILNAGLSYSTYKNDMTRYTSYPEAQGSWTNN
SGALELGGSLALYLPKEAPFFQJYFPFLKFQAVYSRQONFKESGAEARAFDDGDLVNCSI
PVGIRLEKISEDEKNNFEISLAYIDDVRKNPRSRTSLMVSGASWTSLØKNLARQAFLAS
AGSHLTLSPHVELSGEAAYELRGSAHIYNVDCGLRYSF 497602 500415

AGSHLTLSPHVELSGEAAYELRUSAHIYWULGURI SE

CPD_6447 500541 503351

ppD_9-P0 Lymorphic Outer Membrane Procein

FVRPPIAL/MKSSLIMFLISSSLALPLISLINFSAFAAAVEINLGPTNSFSGPGTYTPPAQT

THACTO I YNLTGOVSITNAGSPTALTASCEKETTOHLSFOGHGYOFLUNI DAGANCTFT

NTAAHIY LLSFSGFGYLSI LQTTNATTGTGA IKSTGAGSI GONYSU YRGONFSNDNGGALQ

003 DELGLAPPILTFAKNAATOKGGALYGTYGI TIRHTLINGASFSAFNTANNGGA I YTEA

GOSHALBET INBUTATSARTAJA I YGGSTGAPKFYLT LODM/ELNF LONTA I TEGGA I

YTENDI ZIGAZ REPILTERNINGA I DPAADLOGA I A I ADSOSLALIGI ALFOGHALKOGA I YTENDA I GOTTA I BERNINGA I DPAADLOGA I A I ADSOSLALIGI ALFOGHALKOGA I TEGGATI

YTENDI ZIGAZ REPILTERNINGA I DPAADLOGA I A I ADSOSLALIGI ALFOGHALKOGA I TEGGATI

YZETI ZYFICKEKLISHAFAARAANLIST LOOPLITIACKYOLISLIK/SYTLVAKSFSGSPGGTLL

MOATTFILFTAGS I TIRHINJUN INDELKERKKATIKATOASOFTITLGASIJALUDPSGMYYED

ZOMRHIG/JESCUTLTAANDI AN I HITOLAADPLEKHE I HWYZY SINALSWOEDFATKSKAA

TILMFYRT YNINGEPRICTUNANTIAKOSFVOVRI I COLIVATY/ROOGJETEG I UNGERSKOTE I IKGERILI SA RYVUGATTTI ASSONLI LYAAPCOLIFIKROHEH INKNRAGAYAASL

ILGIILATI SODTALIKYI LASESEQIVLEPAA COLIFIKAMINTYYTOARKGEISWINDOC

AGELAGII JUHTALLIIBE ISHAYEPE FIKVALOV I GIGGJEFREPTLINGE FORGOLI INVOVE

IG PPEEPERIORIKASYEATVI YVADAYYKKHEN TYALQARHITSWKTD TINLSRQAG I GRA

SYNADLANKFOF GIFYAFSPNLEVTSHLSME

CPn_0448 594776 503698
*yxjG_Bs_2 Hypothefical Protein
FIOPSRREIHEWKCILLSSLUMEMMSFFQOPEQCHFDVVGSFLRPESLTRARSDFEEGR
IVVEDMRVVEDAAIPNLIKKOTEACLIFFTDGEFRRYSWDFDFMWGFHGVDRRRDSNDPE
IGVYLKDKISVSKHPFIENFEFVKTFEKGNAKAKCTIPSPSQFFHEMIFAPNLKNTRKFY PTMOELIDDIVFY/POVYQDLYAACCRILQLDCAWCRLLDIRAFSWY/VDSHDRLQEIR POW MINIMOMILAFELLEWILIPAFORYQAFFURFANYO LEEGLEARTOLDCHINANA LEEGLARABELAYOOJERFURLUULUUNIILA TEEFEANYO REFERAA WALAGERLDLAFC CGFASCEGDHRMTEEZOWKK IAFVKEIAKEING

CPN_0449 507231 505330

pmp_10-PMP_10 (Frame-shift with 0451)

EAYTOFROGGISSINNIVQCTTAGNGGAISILAAGECSLSAEAGDITFNGNAIVATTPO
TTKRNSIDIGEST&X:TNLRAISGHSIFFYDPITANTAADSTDTLAILINKADAGNSTDYSGS
IVFSGEKLSEDEÄK/ADALISTIKQDVYLTAGNIVLKRGYULDTKGFTOTAGSSVIMAG
TTLKASTEEVTITGLSIPVDSLGEGKKWIAASAASKNVALSGPILLLDNQGNAYENHDL
GKTQDFSFVQLSALGTATTITDYAAVPTVATPTHYGYQGTWGHTWVDDTASTPKTKTATLA
WINNTGYLENPFRQGFLVENSLWGSFSDIQAIGGVIERSALTLCSDRGFWAAGVANFLKOK
KKGERRKYRHKSGG/AIGGAAQTCSENLISFAFCQLFGSDKDFLVAKNHTDTYAGAFYIQ
HITECSGFIGCLLEKLPGSWSHKPLVLEGQLAYSHVSNDLKTKYTAYPEVKGSWGNNAFN
MHLGASSHSYPEYLLCFDTYAPYIKLNLTYIRQDSFSEKGTEGRSFDDSNLFNLSLPIGV
KFEKFSDCJDFSYDLTLSYVPDLIRNDFKCTTALVISGASWETYANNLARQALQVRAGSH
VAFSPMFBVLQGFVFEVRGSSRIYNVDLGGKFQF

CPn_0490 508121 507180
pmp_10-Polymorphic Outer Membrane Protein
scPMK8-OFSWLVLSSTLACFTSCSTVFAATAENIGPSDSFDGSTNTGTYTPKNTTTGIDY
TLTGDTLUNLGDSAALTKGCFSDTTESLSFAGKCYSLSFLNIKSSAEGAALSVTTDKNL sltg/ssltflaapssvittpsgkgavkcggdltfdnngtilfkodyceengga istrnl slkystgsisfegnkssatgkkggaicatgtvditnntaptlfsnniaeaagga instgn CTI/GNTSLVFSENSVTATAGNGGALSGDADVTISGNQSVTFSGNQAVANGGAIYAKKLT

CPL_0451 508158 511058
php_10-PMP_10 (Frame-shift with 0451)
ATORVKIKILDSCFVIFNLIYLFCFYIDANSSLKNKSITMKTSIPWVLVSSVLAFSCHLO
SLANEELLSPDDSFW3NIDSCFFFPKTSATTYSLTGDVFFYEPGKGTPLSDSCFKOTTDN SLANEELLSPDDSFNGNIDSGTFTPKTSATTYSLTGDVFFYEPGKGTPLSDSCFKQTTDN
LTFLINGHSLTFGFIDAGTHAGAAASTTANKNLTFSGFSLLSFDSSPSTTVTTGGTLSS
AGGVNLENIRKLVVAGNTSTADGGAIKGASFLLTGTSGDAIFSNNSSSTKGGAIATTAGGAIRASCA
RIANNTGYVRFLSNIASTSGGAIHAKKLALSSGGFTEFLRNNVSSATFKGGAIATTAGA
RIANNTGYVRFLSNIASTSGGAIHAKKLALSSGGFTEFLRNNVSSATFKGGAISTBAG
ELISISNRTGNIFFVRNTLTTTGSTDTPKRNAINIGSNKFTEFLRAAKNTHFFYDPISIDASC
ELSLSAEGNIFFVRNTLTTTGSTDTPKRNAINIGSNKFTEFLRAAKNTHFFYDPISIDASC
GTSSDVLKINNGSAGALNPYQGTILFSGETLTADELKVADNLKSSFTQPVSLSGGKLLLQ
KGVTLESTSFSQEAGSLLGMDSGTTLSTTAGSITITNIGINVDSLGLKQPVSLTAKGASN
KVIVSGKLULDIEGNIYESHMFSHDQLFSLLKITVDADVDTNVDISSLIPPVPAEDPASE
YGFQGQWNVNTTDTATNTKEATATWTKTGFVPSPERKSALVCNTLWGVFTDIRSLQQLV
EIGATGMEHKQGFWSSMTNILHKTGDENRKGFRITSGGYVIGGSAHTFKDDLFFTAFCH
LFARDKDCFIAHNNSFTYGGTLFFKHSHTLQPQNYLRLGRAKFSESAIKFPREIPLALD
VQVSFSHSDNRMETHYTSLPESEGSWSNECIAGGIGLDLPFVLSNPHPLFKTFIPQMKVE
MYYVSQNSFFESSSDGRGFSIGRLLNLSIPVGAKFVQGDIGDSYTYDLSGFFVSDVYRNN
POSTATLVMSPDSWKIRGGNLSRQAFLLRGSNNYVYNSNCELFGHYAMELRGSSRNYNVD
VGTKLRF VCTKLRF

CPn_0452 511304 512860
pmp_12-Polymorphic Outer Membrane Protein (truncated)
pmp_12-Polymorphic Outer Membrane Protein (truncated)
pneETMT1LRNFLTCSALFLALPAAAQVVYLHESDGYNGAINNKSLEPKITCYPEGTSYI
FLDDVRISNVKHDQEDAGYFINRSONLFFMGNRCNFTFHNLMTEGFCAAISNRVGDTTLT
LSNFSYLAFTSAPLLPQGQAIYSLGSVMIENSEEVTFCGNYSSWSGAAIYTPYLLGSKA
SRPSVNLSGNRYLVFRDNWSGYGGAISTHNLTLTTRGFSCFENNHAYHDVNSNGGAIAI
APGGSISISVKSGGLIFKGNTASQDGNTIHNSIHLQSGAQFKNLRAVSESVYFYPDHSI
SESKKITDLVINAFEGKETYEDTISFSGLCLDDHEVCAENLTSTILQDVTLAGGTLSLSD
GVTLQLHSFKQEASTLITMSPGTTLLCSGDARVQNLHILIEDTDNFVPVRIRAEDKDALV
SLEKLKVAFEAYWSVYDFPQFKEAFTIFLLELLGFSFDSLLLGETTLERTQVTTENDAVR
GFWSLSWEEYPPSLDKDRRITPTKKTVFLTWNPEITSTP

GFWSLSWEEYPPSLDKDRRITPIKKIVELIWNFELISIF

CPn_0453 513156 516152

pmp_13 -Polymorphic Outer Membrane Protein

NCVLLYLFFYSLSLICRIIWFHLYVOMKTSIKKFLISTTLAPCFASTAFTVEVIMPSENF
DGSGKIFPYTTLSDPRGTLLIFSGDLYIANLDNAISRTSSSCFSNRAGALQILGKGGVF

SFLNIRSSADGAAISVITONPELCPLSFSGFSOMIFDNCESLTSDTSASNVIPHASAIY

ATTPMLFTNNDSILFQYNRSAGFGAAIRGTSITIENTKKSLLFNONGSISNGGALTGSAA
INLINNSAPVIFSTNATGIYGGAIYLTGGSMLTSGNLGGVLFVNNSSRSGGAIYANGNYT
SNNSDLTPQNNTASPONSLPAPTPPPTPPATPLLGYYGGAIYCTPPATPPPPTGVSLTIS
GENSVTFLENIASEXGGALYGKKISIDSNKSTIFLGNTAGKGGAIAIPESGELSLSANGC
DILFNKNLSITSGTFTRNSIHFGKDAKFATLGATQGYTLYFYDPITSDLSSASAAATVV
VNPKASADGAYSGTIVFSGETLTATEAATPANATSTLNOKLELEGGTLALRNGATLNYNN
FTODEKSVVIMDASTTLATTNGANNTDGAITLNKLVINLDSLDGTKAAVVNVQSTNGALT
ISGTLGLVKNSGOVINNGMFNKDLOQVPILELKATSNTVTTTDFSLGTNGXOGSPGVOG
GTWEFT IDTTTHTTTGNKKTTQYLPHPERLAPLIPNSLWANVIDLRAVSQASAADGEDVP
GKOLSITGITNFFHANHTGDARSYRHMGGVLINTYTRITPDBALSTGCOLFTKSKDYL
VGHGHSNVYFATVYSNITKSLFGSSRFFSGGTSRVTYSRSNEKVKTSTTKLPRGRCSWNN
NCWLGELEGNLPITLSSRILNKGIIPFVKAEVAAVATHGGIQENTPEGRIFGHGHLLNVA
VPVGVGPGKOKNSHNEPDYTTIVAYAPDVYRHNPDCDTTLPINGATWTSIGNNLTRSTLLV
QASSHTSVNDVLEIFGHCGCDIRRTSROYTLDIGSKLRF

CPn_0454 516179 519115
pmp_14 -Polymorphic duter Membrane Protein
GMPLDFKSGSFCLLACLGSASCAFAETRLGGHFYPPITHYGEEILLTSDFYCHNFLGASF
SGSFINGSONLGLLSKGLSETFTSGOAPPNGHYALLSAAETLTFKNESSINFTUNGCTGL
GGLIYGKBLYPOSIADLIFTTNRVAYSBASVTTSATPAITTYTTGASALQHTDHJTVENI
SGCIFFGRINANESASIAGSSFTAVKFINNTATMSFSHINFTSSGGGHYGGGGLLEFING
GGLIFTANEGANS GGYTPSSGTYALGGGGAICIPTGTFELKNNGGKCTFGYNTTPNDAG CCLLETANECONES GOVTESSCTYALGEZGALC LPTGTFELKNNGGRCTESEM TEPHDAG ALVAETENIN VONCALLLDENTAANGKALGEZGALC LPTGTFELKNNGGRCTESEM TEPHDAG ALVAETENIN VONCALLLDENTAANGKALGEZGALT LTGT EVERGEFELKEN LTGTEN ALVERGESKE LEIGHT AFGENMUNTER LENALTVEAGGE LEIGHT LLGEN DE LEIGHT LIGHT LEIGHT LE YGGYQAALGMATTDHTTLGLGFQQLYGKTNANPYDSR LLGSFFGQFPIVTQKSEA LLGWKAAYGYGKNHLMTT/LRPDKAPKSGGOMINNSY CEMPFLMCLLTRPLAQA WOLGGFISAEFLGOMGSKFTETGDLQRGFSRGKGYNVS CSSQWFTPFKKAPSTLTI KLAYKPDIYRVNPHNITUTVVSNQESTSISGANLRRHGLFVQIHDVVDLTEDTQAFLNYTF DCKNGFTNHRVJTGLKSTF

CPn_0455 520363 519458
No robust homolog present in Genebank/EMBL as of 11/7/98
PROPRIETO PROFESSIONAL GENERAL CONTROL OF THE PROFESSIONAL ASCRIVATION FROM THE PROFESSIONAL ASCRIVATION FOR THE PROFESSIONAL ASCRIVATION FOR THE PROFESSIONAL ASCRIVATION FOR THE PROFESSIONAL TRANSPORT OF THE PROFESSIONAL ASCRIVATION FOR THE PROFESSIONAL TRANSPORT OF THE PROFESSIONAL T

CPn_0456 521568 520327
No robust homolog present in Genebank/EMBL as of 11/7/98
IPCTFESKRKFLMTHCLHGWFSVVRHHFVQASNFSRPLYSRITHFALGVIKAIPIVGHLV
MGVWLISHCFERGVSHPGFPSDIAPILKVEKIAGRDHISRIENQLKSLRRTIEVEDLDK
VHGQYQENYADMASSEVLKLDKGVHVSELGRAFSRVRNRITRSYSYAPTPQLDSIAIVG
IDLVSPEEQENLVRLANEVIQLYPKSKTTLYLLIDFNKEWVGDISSDKEKQLRSLGLHSE
VQCLSVLEPQGASGEDTHKFDLMVGCYGKDSYLREGKILQQALGTSIGTIFWNVMHTLP
SRYRSRLSLPINTEKDKTELYKEISRTHHQLHTLGMGLGAQDSGLLLDRGRLHAPLSQGS
HCHSYLADLTHEELKILLFSAFVDAKNISKKELREVSLNFANDTSVECGCAFYF

CPn_0457 523886 522120

No robust homolog present in Genebank/EMBL as of 11/7/98

VFLPSRVMASCLSAWFSIVREHFYRAFDFSLFFCARITEFVLGVIKGIPVVGHIIVGIEW

LVSRYLESFVTKPFFVSDVVSLLKTEKVAGROHIARVVETLKRQRVAVAPEDEDKVHGKI

PVHPFGGIOPVEVLTLYPEVQDATIGLAFSKIRNRVRQAYLQAPRPKLQKIYIIGNDMNP

FEVDDFLHLARLCNETQRLYPDATISLYLTASGGRNAMDKKNRKLLSDCELNPKIACLDF

NQGDVVKQATCDCWMVYHGENDCGTLUD1GELEKSGEFTPHIHVGQKPLGSLMDFSPF

SSLDMKGDKEKALEYSELEKEQLYSRLVYVGERSSVLSLGFGDSRSGILMDFKRVHAPLS

EGHYCHSYLADLENGLQKTILAAFLNPKELSSTILQPISLMLILNSKTYLRQHFGFFER

MSRSDRNVVVVVCDSWAGTDWKEEPSFOHFIMELECRGYSHRNIFAFRSNSWCVEERRIL

NESSQEKAFTMIFCEDSVSQGDIRCLHLASEGMLCGKECYAVDVYTSGCANFMMEEVLTL

ERESNLWNRKHGLWKREVVKQKQEAALDQDESEIYVCNQLTAQQNFACS

CPn_0458 526344 524236

No robust homolog present in Genebank/EMBL as of 11/7/98

PEECYLKLFYSNFIFFVVMPIFYISSWISTVRQMFVKAFDFSRFFCSRVINFALGVIKA
IPEVGHIVMGMEWLVSSCVAGIITRSSFTSDVVQIVKTEKALGRDHISRVAEILQRERGT
IPENQNKVMGKFPVCPFGRLKSEETLKLKPGERGTLDTVFSPIRTRVTRAYLQAPRPE
IRTISIVGSKLKTPODFSOFVSLANETORLHPEALVCLYLTGLNRESGMCDTTTAEKKQY
LHRSGLDSRIOCKDSKEDDAGSPEDREJMIGTYSREQOMNIOGOVIQOCLGKSADPIFWI
HYTEDTKDFYYPPNFTSYSHTRQSTDPTSPPRLPESEGDKDSLYGQUSRSYHHEYMLGLG
LKPEDDAGLLMDPDRIYAPLSGGHYCHSYLADIENEDLHTLVLSFFLDFGNLSSEDLRPVA
FNIÄRHPLELDSLFFRLVAGQODGSNIVTLAHGTRPEDLDDFSNILTRRLQMSGYSYL
NIÄSYKSRKMIVKERQFFGDRSEGKSFTLILFEDPISAADFRCLQLAAEGMVAKDLFSVA
DICASGCSCIOFSEMOSPQAIEYRGMEARVEDEAGEEAREPVIYSQDQLSSMLTTQONFV
FSÜDAVKOAINFRFSKGKLITMERKALGEFFLTATFSYLGSGERNENMGKRTTEEHEVVI
SEEELDRMVQVLPAEVPADSGNDPTRPVPNPDSNPDSSQNEGS

CET-0459 527062 526619
NOT-cobust homolog present in Genebank/EMBL as of 11/7/98
STEXIOMHPGLERWRTSTINKLREEGSVSFREYFRAYMCDKIVAQKNFLFTLDAVIKQAGWR
SQEKLNLFYVESQALGREIKVSLEFYIQSMVGILGSQRTKKSFKFSVDFTPLEQALQERC
SSEDDEDATATSTATGATASPTDMHEDE

CHE-0460 527840 526992
No. robust homolog present in Genebank/EMBL as of 11/7/98
VEHILNFALEETPSISVYYQEQEKLSPCDHSPEIGKKKRWNKLESFSTYCSLFMSVKDM
YKANLGIONSLSGWLLDPYRVCAPLSSPYSCPSYLLDLONKELRRSLLSTFLDPKNLTSF
TFRSVSINFGNSSFQGRWSEFLSRVLHDEKEKHVAVVCNDAKLLEEGLSPEALSLLEEDL
RESGYSYLNILSVSPEGVSKVQERQILRRDLQGRSFTVMITDLPLGSEDIRSLQLASPRI
LVSSSLDAADACASGCKVLVYENPNASWAQELENFYKQVERRR

CPHT-0461 528647 527844
No robust homolog present in Genebank/EMBL as of 11/7/8
ISIVACPSISSWFTVVRQHFVNAFDFTHFVCSRITNFALGIIKAIPVLGHIVMYIEWLIS
WIPRHTVRHGMFTSDVSSAIKVEQTRGHNCLAPLEAYLSSLRVPISQEDLGKV/GGTFED
PFVDITPTEIVQLLPDEELSTVDEALQGVRSRLTYAYRSVEKPMIQDLALVG/GGLRDSAD
LINFVRLANGVQNHYPHTKVKLYLAKNLADVWDCEISEEEKGQLRALGLDPXIESISLTS
AGLPSVPEVATVDFMITCYGKDQEVQDP

CPn_0462 531124 529037
No robust homolog present in Genebank/EMBL as of 11/7/98
LIFYLFLNLYIACVRFHFQCWFDPMACYISIWISTVKQHFIRAFDFTRFLGSRITNFALG
VIKAIPILGCVVIGVSWLVSTCSARRFGKPAFTSDVASIVKIEKTRG/NPLAWVEQYLRQ
LRVRLPEDDLGKHHGKVSRDYVCDRTPQENLMVPHQYLGELGRAFYGIRNRVTKAYQRV
TPLEVPCLTLVGFDILDPEDQVNFVRLANGIQTQYPQTQIKLYLLISIOKIMNQCDGTISQ
EKEQQLRSLGLDAKIKCVSAPALLLQKYLQSENLPSCDLLINYYGKQOSVRDVDSIKSLL
NLSSEHIPAISVTYRPDDFFYSYFFFPGSQGTAPDQRIPMSEQFHLGTYTTLSNPRCDR
YAVHLGKBEDPASGVFLDPLRVSAPLSGEYSCPSYLLDLKSEELFFLLSAFIDPNNSGQG
NPRPMSINFGNSPLGQRWSEFLSRVLHDETEKHVAVVCNNPQL/KKSFPSHSLSLLENEL
EEGGYSYLNIVSVSQERTCVKERRILSSDPSGRSTTVILTDL/EGSSDIRNLQLASDRIL
VGSALDAADACASGCKILEYEDPEDEDWAQOYASFYRNIDRAGUDQRQGIPGEPLGVSATS
RVVLEKDIVFNLNAVIQ\AMWKFKKRDLFAVESQALGDMGRALEGYIGSSLLVEGTIQP
QVACNNNVSPATLDEAVCAACDSAQDAPSEENNTDD

CPG_0463 532480 531191

TO TOBUST HOMOTOR PROSENT IN GENEDARK/EMBL as Of 11/7/98

HIGHPERMEDILGTENCRIPPRINTSTYG [PIDETCHAFVDSMMKKYNGQDAKELYTFLSR

GHERZOLGLWEGLERELIFLEPEKMLCAPLSEDHYCHSYLVDLVDDHLKDLILSMFLDPO

HIGHELIKVGINVGDISEPOLOGKOPLEMVLRDETY ZNVVVVYRKVULSLPATOVCKLYEL

HIGHETAVATARCKILKI DHTMREGTESGRY TV TCALYLGTOMRSLOLASERIM

VOMERULVDAYAARCKILKI DHTMREGTESGRIADFADAVDVSAGRISREFKLITOANQG

HIGHETALPEKTHERGELAFCDRYTVTHHEI FURLDAAT KOAVWTHKHESLI DKECEALD

LPTOCHESIVSTLEYVTNEHEKTSKGPFIQKELIADCOPLKEALFPGSDEDVPSTSEDPS

DDHEGDLEDD

CPn_0454 512.366
No robust homotog plane in Genebink/EMBL is of 11/77/98
SLETRGRFTEICLGLEFEIGEKELLDEGESTALNLERIFAPLENRYTTEYSRARQPDL
HRIAIVYIGVLDSESSKILERLISYMSC/YSESOMYLEFFMCKNYNGSAVLSKLHVENLH
IRCGFFSEDAVPESEPFDLSIYVHTDRSCPLPTKKRSSSWELCTVELPESIYPQSEFLLM
RDBM/S

CPn_0465 533278 512871

The test-sum boundary process of the desired bank MMRL in of 1177.08

THE TEST-SUM PROCESS OF MATERIAL AND THE CARDIFFECT CONTROLLING STREET AND TRANSPORTED TO THE TEST OF TH

CPn_0466 533/18 536537

pmp_15-Polymorphic Oyter Membrane Protein
TSMRFFCRGMLIPFTFVLANGGLOLPLETYTILSPEYQAAPQVGFTHNQNQDLAIVGNHN
DFILDYKYRSNGGALTKKILISENIGNVFFEKNVCPNSGGAIYAAQNCTISKNQNYAF
TTNLVSDNPTATAGSLLGGALFAINCSITNNLGQOTFVDNLALNKGGALYTETNLSIKDN
KGPIIIKQNRALNSDSLGGIYSGNSLNIEGNSGAIQITSNSSGSGGGIFSTQTLTISSN
KKLIEISENSAFANNYGNFNPGGGGLTTTFCTILINNEGVLFNNNQSGNAGAHAKSI
IIKENGPVYFLNNTATKGGALLNLSAGSGNGSFILSADNGDIIFNNNTASKHALNPYRN
AIHSTPNNNLQIGARGYRVLFYDPIEHELPSSFPILFNFETHGTTGTLTSGEHVNONFT
DEMNFFSYLRNTSELRGGVLAVEGAGLACYKFQRGGTLLLGGAVITTAGTIPTPSST
PTTVGSTITLHNIADLPSILSFQAQAFKINIYPTKTGSTYTEDSNPTITISGTLTLRNS
NNEDPYDSLDLSHYEKVPLLYIVDVAAQKINSSQLDLSTLNSGEHYGYGGIMSTTWVET
TITNPTSLLGANTKHKLLYANMSPLGYRPHPERRGEFITNALWGSAYTALAGLHSLSSW
DEEKGHAASLQGIGLUVQKKNAFKGFRSHMTGYSATTEATSSGSPNFSLGFAGFFSKA
KEHESONSTSSHHYEGKCIENTLFKENIRLSVSLAYMFTSETHTIMTMGCLLEGNSGGSF
HNHTLAGALSVFLPQPHGESLQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVS
LFVGIRASWKMHRVPLVWLTEISYRSTLYRODPELHSKLLISQTWTTQATPVTYNALG
IKVKNTMOVF PKVTLSLDYSADISSSTLSHYLNVASRMFF

CPn_0467
536528 539434
pmp_16-Polymorphic Outer Membrane Protein
NEILTISDONRKIKEPLVSKTPPKFLFYLGNFTACMFGMTPAVYSLOTDSLEKFALERDE
EFRISSPELLDSLSTLTGFSPJTTFFGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCN
NFLLSWEDHAFFSKNIAJGTGGATACOGACTITKNRGPLIFFSNRGLNNASTGGFTRGG
AIACMGDFTISONOGFFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFNNTAPSGGGAL
RSDYTTISDNTRPIYFKNNCGNNGGAIGTSTVYAKKNNGSGYIFNNNTALSGSINSGNGS
GGAYTYTNLSIDNNPGTILFNNNYCIRGGAICTOFLTIKNSGHVYFTNNGNNGGALML
LODSTCLLFAECONIAFONNEWFLTTFGRYNAIHCTPNSNLQLGANKGYTTAFFDPIEHO
HPTTNPLIFNPNANHGGTILFSSAVIPEASDYENFISSSKNTSELRNGVISIEDRAGWQ
FYKFTCKGGILKLGHAASIATTANSETPSTSVGSOVIINNLAINLPSILAKGKAPTLWIR
PLOSSAPFTEDNNPTITLSGPLTLLINEENRPYDSIDLSEPLONIHLLSLSDVTARHINT
DNFHPESLNATEHYGYGGIWSPYWETITTTNNASIETANLLYRALYAMMTPLGYKVNPE
YOGDLATTPLWGSFHTMFSLLRSYNRTGBSIERFFLEIGGIADGLFYHONSIFGARGFR
IQSTGYSLOASSETSLHGKISLGFAGFFTRTKEIGSSNNVSAHNTVSSLYVELMFFORAF
ATSTVLAYGYGGHHLHSLHPSNGCOAEGTCYSHTLAAAIGCSFPWQKSYLHLSPFVQAI
AXRSHOTAFEEIGDNPRKFVSQKPFYNLTLPLGIGGKWGSKFKYPTEWTLELSYOPVLYQ
ONPOIGVTLLASGGSWDILGHNYVRNALGYKVHNQTALFRSLDLFLDYGGSVSSSTSTHH
LOAGSTLKF

∠Pn_0468 539608 540432
pmp_17-Polymorphic Outer Membrane Protein
IYKLLDNKLMIFYDKLYPHIKVMmFMRFICLSILSTALCCSLSGNEVPNLASCOMSRKDI
SAFHTSPSFRLMVTPEPLVSSFRPSNLLNGFGHDITQDITITGNSINSVIDYNYHYEDGG
ILACKNLFISENKGNLSFERNSSHSSGGALVSVRECWISKNQNYSFISNAASLATTTTSG
FGGGIFCRAVNIERNYQNIQINDNASGGGVYYFLP

CPn_0469 540399 541460
pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0469)
CFRTRGGIFSALGVIISSNKEIIEISNHSASSINTASGKLYPGGGIMCTSLVIENNPKG
LIFNNKTAALSGGAIHTRSFIFONNSPTAFINNSATSGGALINLSGIGSTPONFFLSADY
GDILFNNNTTITSSSPOPGYRNALYAAPGINLKLGARGGYKLFYDPIOHDOTTTDPIVFN
YEPHHLGTVLFSGINVDSNATNPLNFLSKFSNSSRLERGVLAIEDRAAISCKTLSOTGGI
LRIGNAALIRTKGPGSSINFNAIAINLPSILOSEASAPKFWIYPTLTGSTYSEDTSSTIT
LSGPLFFLNDENENPYDSLDLSEPRKDIPPELPRCDCKKNRYFESHCRSHELR

CPn_0470 541357 542532
pmp_17-Polymorphic Outer Membrane Protein (Frame-shift with 0470)
ISLNLERISPLLYLLDVTAKKIDTSNLIVEAMNLDEHYGYQGIWSPYWMETTTTTSSTVP EQTHTNHRQLYVDWTPVGYRPHPEHRGEFIANTLWQSAYNALLGIRILPPONLKEHDLEA SLQGLGLLINQHNREGRKGFRNHTTGYAATTSAKTAARHSFSLGFAMFSKTRERGSPST TSSHNYFAGLRFDSLLFRDFISTGLSLGYSYGDHHMLCHYTEILKGSSKAFFNNHTLVAS LDCTFLPARITRTLELQPFISAIALRCSQASFQETGDHIRKFHPKHPLTDLSSPIGFRSE WKTSHHIPMLWTTEISYVPTLIYKRWPEMFTTLLISNGTWTTOATPVSYNSVAAKIKNTSQ LFSRVTLSLDYSAQVSSSTVGQYLKAESHCTF

CPn_0471 542561 545401
pmp_18-POLymorphic Outer Membrane Protein
TYONNRSLKSSFFYGALILGKTTILLMATPLSDYFDMQANQLITTLFPLIDTLINMTPYS
HRATLEGVRDDTNQOI VLDHONS IESWEENFSQDCGALSCKSLAITNTKNQILFLNSFAI
KRAGAMYVNONFDLSENHGS II FSGNLSF PNASNFADTCTGGAVLC SKNVTISKNQTTAY
FINNKAKSSGA TQAALINI KINKDNTDGPLFFNNANGATAGACHANAGRIENNSOGTYYLI
NOSGLGGA IRVHQEC ILTKNTGSVIFNNFAMEAD ISANHSSGGAIYC ISCSIKDNGIA
AFDNNTAARDCGALCTOSLT IQDOSPVYTTNNQTHWGAI MLRODGACTLFADQDII IFY
NNHHEKDTESHIVSVNCTRINNSLTTVGAGQGIIGATEYDD ILDRYT IQNS IQKFNRNPEHLG
TILFSSTY I PDTSTSKDDF ISHERNH IGLYNGTLALEDRAEMKVYKFQCGCTLRLGSRA
VFCTTDEEQSSGRVSVIN I NNLAINLEG ILJNRVAPKLWIR PTGSAPYGEDNP I INL
SGFEGALLDDENLDFYDTADLAQFILEVFLLYLLDVTAKHI INTDMFY PEGLAPTOHYGYQ
VWGFWI I ET ITTSTDSSEDTNTTLIRGLYGMKTPTGYYVNFENKOI IALGAFWGSFHNLE
ATERYGTGAG I TATTAJFACHLEFJIGENNDAKSFHNEATGYGJYTTSTNAINHSEYN
FOGLEDNLYESBENSVAGHTTTALG INDIWLOERFSTSASLAYSYSNHHI IKAGYYGK
KOPLYNTOFI GYSWEGKERLFFWHI IELAYYGVLYQNDENNGLEGGEKARKESVH
KEPLYNLTVELG GYSWEGKERLFFWHI IELAYYGVLYQNDENNGLEGGEKARKESVH
KEPLYNLTVELG GYSWEGKERLFFWHI IELAYYGVLYQNDEI KHYGLEGGEGGEKARKESVH
LANALAFKGROQ I F I FEKLEVFLLYY OUVGSTTTHIYLHAGTTFKF

CFn_0472 547914 545581

CPn_0473 549602 548070

No robust homolog present in Genebank/EMBL as of 11/7/98
GSIMAVGGVGGSRSPSPIPPNRRNSEDGKVSPKDNLGEHTVSSDSSLASGGPTIEERKA
QLGGTDKIPLPSVKEPGDSQTSGRSGVLQRIWKGVKGVFKKTPQARPEVSSPRLPSHVQH
GQRLPGLEGFRDRIQKRSENPEADLGKMKRSYSDGDLDRVGHDSNEDSTEDSRSEGGEPS
SKSSSFLSGVRGAVSKVHGALGDIKGKFQRSASEDDLTTQGEDSAGDTVKERRSEEAEAS
SKSSSFLSGVRGATSTVQGALGDAKEKVSAFGEQAAGAIRSAPGNIRTRFQRSSEGGLS
VKNKAAKHLRKALENLEKVAPEQVSPEVASRVQSLLARMEQLTHQEPTVEDLITFVESN
VGSDSVEYASIVPQDGSQAPAETAEAPETGGVEGSAAQGAMKALRDFVVSIFQAVASFFR
AIASRLSSARRESAVDDLASSSNTOWFVEQEGVSNPSAAPSLSFAEEIARRAAEMSNRNA
QSLEKLESGNVTDPVIQQGLGLARSFAPEGQ

CPn_0474 551600 549807
CT365 hypothetical protein
LKIIISISFMSTSPISNDPRYLSLSNATEKTSLLANSRSLSPVPNSLVPSNPEDTGLRKS
IFTHSVTLFAGLVVLLVAVSVVVVALTVLAPGVPQAILLGIAISGVGIGGFSIMKSLVYM
VRDVMSPRMGESSRIKSALAVGTGFTVMGLVMKVGANFVPGGYGGLVGSLGSSAYSRGSQ
TTLASFSHYIYTKFFRSEKVAKGEKLTEAETIKEAKKLHYITLSIATIGVGLAVIGILLA
LGGTVLLGGAPATIAIILAPPLISIGLTTVLQTILHSSIGKWRAFLLTQEKKDLFVDTSL
KDIRLEKLPPSEVEESETSGSVIEVPDSEGIAETRISAEEIDTRLSLTTRGKVIFALATL
LLLASIAAFIVTGFGGLTVMQVLLVASVGSAVASVTLPMVSSGFSVYAYQLKARINISKL
LLLASIAAFIVTGFGGLTVMQVLLVASVGSAVASVTLPMVSSGFSVYAYQLKARINISKL
RWKEAKNKKRVROFLIESGVIASDREFNOMMTVVKKQIQKTDAAIREEVRNFEKGGEVN
SALVGGILLGVGTGIMLLALVPAFAPIVPGILALGGSTLGAGSILMRKFVNMLYDELVK
LYERRNRRELLYGPESKMRSIATDLVVEALAASHDHLFDLDGPVDFIDVDVDIDGAA

CPn_0475 553850 551685
glgB_Glucan Branching Enzyme
PSWDKLIHPMDLDLLVSGRKDPHKLLGILASEDSSDHIVIFRPGAHTVAIELLGELHH
AVATRSCLFFLSVPKGIGHGDYRVYHQNGLLAHDPYAFPPLWGEIDSFLFHRGTHYRIYE
RMGBIPMEVQGISGVLFVLWAPHAQRVSVVGDFNFWHGLVNPLRKISDQGIWELFVPGLG
EGGEVKHUVTOSGNVIVXTDPYGKSFDPPPQGTARVADSESYSWSDHRWMERRSKQSEG
PVTLYEVHLGSWQWQEGRPLSYSEMAHRLASYCKEMHTHVELLPITEHPLNESWGYQVT
GYABPTSRYGTLQEFQYFVDYHKENIGIILDWPGHFPVDAFALASFDGEPLYEYTGHS
OALHPHWNTFTFDYSRHEVTNFLLGSALFWLDKWHIDGLRVDAVASKLYRDYGREDGEWT
PNAGKENLESIEFLKHLNSVIHKEFSGVLTFAEESTAFFGVTTDVDCGGLGFDYXMNL
GWMHDTFHYFMKDPMYRKYHQKDLTFSLMYAFQESFILPLSHDEVVHGKGSLVMKLPGDT
WTFFAQMRVLLSYGICLDKKLLFMGGEFGQYGEWSPDRPLDWELLNHHYKKTLRNCVSA
LNASYIHQPYLMMQESQECFHWVDFHDIENNVIAYYRPAGSNRSSALLCVHHFSASTFF
SYVERCEGVKHCELLLNTDDESFGGSGKGNRAPVVCQDQGVAWGLDIELPPLATVIYLVT

CPE_0476 554877 553858
CT865 hypothetical protein
GRGMRADMGDCMIDINGHFRPTMVPGQKLPIPGSLLYAQVFPTLWRLFSSKHEILMEQT
LQVQSPLKRFAVFQDLHRGGLAVTSERYYKYYLLPSGECTQSIKGKLPSAAQAGPLISLGV
HKBREWGKVRCRDLKELIPLWFRFAMAPKGSYRDLETTAIGSILVKTAHQRUHAETTE
IAPALLSIALAGFSECFLPRSYDEEFGGILPQDGDPEGGVPFELLSYSFGMIODIFLAHQ
GGEVEILPALPPEFFCGRIHVALPNLGTLSIVWTKKTIRQVELHAEYSGEVFKFCSSL
CSARLREWSERRLSGSKRLSLGETLEIKAGTTYLMDCFHK

CP5_E477 556112 554844

*ydeV_Bs Hypothetical Protein

RYMTVAEVKOTFKLVCLGCRVNQYEVQAYRDQLTILGYQEVLDSEIPAD/CIINTCAVTA
SAESGRHAVRQLCRQNPTAHIVVTGCLGESDKEFFASLDRQCTLVSNYEKSRLIEKIFS
YDTTFPEFKHISFEGKSRAFIKYQDGCNSFCSYCIIPYLRGRSYSFPÆKILAEIAGVVD

QGYREVVIAGINVGDYCDGERSLASLIEQVDRIPGIERIRISSIDPDDITEDLHRAITSS
RHTCPSSHLVLQSGSNSILKRWNRKYSRCDFLDCVEKFRASDRYAFTTDVIVGFPGESD

OPFEDTLRIIEDWGFIKVHSFPFSARRTKAYTFDNQIENQVIYEKKYLAEVARRVGQK
EMMRRLGETTEVLVEKVTGQVATGHSPYFEKVSFPVVGTVAINTIVSVRLDRVEEEGLIG
FIV

CPn_0478 557640 556210
hflx-GTP Binding Protein
whGGPLDTIDTPOEQGSQSFGNSLGARFDLPRKEQDPSOALAVASYQNKTDSQVVEEHLD
ELISLADSCGISVLETRSWILKTPSASTYINVCKLEEIEFILKEFPSIGTLIIDEEITPS
QGRNLEKREGLVVLDRTELILEIFSSRALTAEANIOVOLGARVLLPRLFKLWCHLSROK
SGGSGGFVKGEGEKGIELDRRWYERIHKLSAQLKAV/KQRAERRKVKSRRGIPTFALI
GYTNSGKSTLLNLLTAADTYVEDKLFATLDPKTRKCVLFOGRHVLLTDTVGFIRKLPHTL
VAAFKSTLEAAFHEDVLLHVVDASHPLALEHVQTTYDLFGELKEKPRIITVLNKVDRLP
QGSIPMKLRLLSPLPVLISAKTGEGIQNLLSLMTEIJQEKSLHVTLNFPPTEYGKFTELC
DAGVVASSRYQEDFLVVEAYLPKELQKKFRPFISYFFPEDCGDDEGRGPVLESSFGD

CPn_0479 558434 557616
phnP-Metal Dependent Hydrolase
A IGMYROLOGES IGKLVFLGTGNPEG IPVPFC SCRUGONTG IHRLRSSVLIQYONKTLVI
DAGPDFRTOMLVAGVSELD WELTHPHYDHIGG IDDLRAWY IVTORSLEEVLGAGTYRFL
NKAKEYLFATPRIVESSLPAVLEET ILNEDGSOEFEG IPYTYVGYYOKSCHVTGFRFGNL
AVUTDLGYDAK IFGYLDNVETLILAGAGPYETT I FFG HKGSHLTVEBAKAFANHAG IKN
LI ITHIGHGLEAERDGHPEVTFAYDGMEYLWTL

CEN_0480 559375 58650
CET 88 3 hypothetical protein
CHARLEWITKERLEEFVERKEQDEMISSILECTRAGHTEEYKRRYFYCOLGAEVVSPYVV
PALTVYDVG APPENTLOVLECKOHING DLÉVHED TECLMALEEVERGAFOLEGAMYELGS
QVENED LEG EVENVEGGLET FAGLEVGVMVEAPLITAGLEAWV FEG TICKVGA LLELEFATL
MAYLDE HEVREWLEHLEHEY LTOOKICKO LOAHSQNYDV LTEYPATGALEQP LTKLPNGSER

CPn_0481 560773 550330
NO robust homolog present in Genebank/EMBL as of 11/7/98
SCLRIEGILMATSVPVTSSTS/GEANS/NERTERTSRMYYAALVLGALSCLIFIAMIVI
SPOVGLMAVVLGFALGCLLLSLAIVFAVSGLVLGKTLEPSREATPPEIVACKEMTTOODV
LGNEYWRSELISLFLRGDLHEBLIVTJSKDRSLDIDOSLQNILKLEPLSTTLSLLKKDCVH
INIILHLVRGWNLLGVDLSFEVTAMSEELLLFLIEEQYYSPDILKLIRYGDALQATSPLM
HAMBBLISLFEVTAMIVATORAYERA FEBRUARYERA FEB

23

CPn_0483 561830 564964

No robust homolog present in Genebank/EMBL as of 11/7/98

IILIKKRAIFEMMFPIPPHCPPNNKNNFYHLTTDTKOPLLLÄILATIGYVLLHIITLGL

LLIHYYKHHÄVVRKEGLFYPPTLEKPGPPKNT LEIAKOPPKOEDKKPDVPKRGTPPPED

TPPPPPKAPPASPKVPKOPADKKPTPPPEAPPPVRVATPMPLRPSSGOYWGCLNRMVS

MVLRRAPLPLPAMQVDPILGOFNPHFVASYPNRIDNEPMYFQIKQFKKIAQNPDLPQMWS

MVLRRAPLPLPAMQVDPILGOFNPHFVASYPNRIDNEPMYFQIKQFKKIAQNPDLPQMSR

RLAQISLGOALYLNDNYLVNVPGGGNCFYRRYAVGWLSALVEESSRNDIVFEQEATRLL

DLPFASSPANANLCAEMAELLQLCSTYCSFIDLYDGVILSQKHTATLIAFLRKLSAYAI

RQOIAASSNEETARALFISDMQDDLLPSVLEFLANNPYSELFFONLIONSALPYMGSNR

LYLLLEHLPALFLTDAELQMSPEDQQLKQYEREIREAFAKLSRIADSGWTERFNAI

VKDHMPEAIRCQYSRFLATIENRRSGDLPWSPALSFFAFLCTCPSVRFHKLCATFYKSLE

DII JASAPPQRSIGEILGISNASLSYLMEDLDSSWQREVISSNIMTILTHEELTLESSM

PQLETLHKRIANLLKNVISTSFETPPLSNQPDLLSNLVNKLLVAIHSKLELKEHFNTVCS

ARSLRITRDEGSGLSQEQDLLYTQAVQLIFFILQHPQVNNRPETKDAVKELKMLLLFPL

ZYKFKKVEHEKKLQKLLRSILGSLVLKPPARYPSTPSNKDKETFCKFWSRHPEVMVLDPIL

ZYNCMGFLRATFPNQLETEAILLEKEIESTFRNGMNVFLTTALNLFGSKLGSPSSFTALS

DOFSKSFLIFCFLNNYPKLLGKKTPLARLDAFQREASHRFTQVKDKLLISLKYGFPLAT

ATINGYSRARDQLICNLLKNTVTASDGFCRSGFRQSLIGYLHSLSSNELGDILDDVKEQA

EANDVAAMTTVPLQFFAVCLINSDRDTVSEENIENFVAMHGFLNTISPERDARIFLIRFP

NHYGCLLPRNPRTEDQNSKPDSSNP

CPn_0484 564931 565824
arod-Deoxyheptonate Aldolase
RSELKTQOLKSLVLHEVLILTFTYPLPRILKQHPDEVHTVPISPNLSFGEGSPILIAGPC
RSELKTQOLKSLTVKEAGAQVFRGSIRKPRTSPFSFGOWEKECVLHKEAQSIHGLPTE
TEVLDVRDVEITAEHVDILRIGAKNMHNTPLLQEVSKSHRPIILKRSPAATLEEMLCARE
VILASSPSCPGVILCERGIRTFEHSTRYTLDLNTVALLKEISSLPVIVDPSHAAGKRSLV
LPLASAGLSVGADGLMIEVHAHPEKALCDAKQQITPEELHLFAKKHFCPSESRAHAIS

CPn_0485 565993 566229 CT382.1 hypothetical protein QPIGRIPTRYFLWRFHIKQACKFYLLQCLLCALYWLLKYCRKLLKGTLHHSEETLYQALL SSLIDLLYQLKQLPAPTNE

CPn_0486 567799 566405
hypothetical proline permease
AQHRSLLKGNIFHLECGVLYFMNFSLFLFFLIAIQGICLYVGRRGSKKVEDRESYFLAGR
SLKIFPLMMTFIATQIGGGVLLGAAEEAFCYGYGGILYPLGVALGLIFLGMGPGKRLAEG
SLTTVVSIFEVFYGSKKLRKIAFLLSAGSLFFILVAQVIALDRLFSSFPFGKYVTVAFWI
VLASYTSTGGFRGVVATTDVIQAGFLLIAVLVCGVSVMLSVYBKLSVLDPFOSLFCAKLSN
WIFMFMLFMLVEQDMVQRCVAASSPKRLQWAAVGAGLVLLLFNFIPLFLGSLGAKAGLKA
GCPLIDTTAYFCNPSLAAVMAAAIGVAILSTADSLMNAVSQLTAEEYPTLKAFYYRYLVL
GLAVAAPLVAIGFTNIVDVLILSYSLSVCCLSVPVGFYLLAPKGRRVSGAAAWAGVLVGA
LGYGWQIVSLGMFGELLAWVGSLVAFSFVGFIEITWKNKVKTQT

CPn_0487 569833 568112
CT384 hypothetical protein
RRTGGISLTYSSFRWASFRCYSLIFFCFCGSLFGSESLRYQLLIQDFAKVSEEGIGLLES
KEYSLLQAKLVLRALAQNSSFDDWFRSFKKCQISYPELAHDRDVLEEFGIQVLREGIENP
SVTVRAVSVLAIGLARDFRLVPLLLQSCNDDSAIVRSLALQVAVNYGSESLKKAIVELAR
NDDSIHVRITAYQVVALLQIEELLPFLRERAENKLVDSVERREAWKACLELSSQFLEFGV
AKDDIDQALFTCEVLRNCMLPETTEIFTELLSVEHPEVQESLLLSALAWSHQLQNHKEFL
SKVRHVMCTSPFAKVRFQAAALLHLHGDPLGRDSLVEGLRSPQPLVCEAASAALCSLGIH
GVPLAKEHLESLSSRKAAANLSILLLVSREDIERAGDVIARYLSNPEMCWAIEYFLWDAQ
WNLRGDTFPLYSDMIKREIGRKLIRLLAVARYSQAKAVTATFLSGQAQGWSFFSGMFWE
EGDVKTSEDLVTDACFAAKLEGALASLCQKKQQASLQRVSQLYNDSRWQDKLAILESVAF
SENLDAVPFLLDCCHHEAPSLRSAAAGALFSIFK

CPn_0488 570147 569767
hita-HIT Family Hydrolase
RKLPTCFAVMVTRSRDHMTVFKQIIDGLIDGEKVFENENFIAIKDRFPQAPVHLLIIPKK
PIPRFQDIPGDEMILMAEAJKIVQELAAEFGIADGYRVVINNGAEGGQAVFHLHIHLLGG
RPLGAIA

CPL_0489 57:037 570090
CT387 hypothetical protein
RIVFALENYFOLVYINEDMURRIVOMOTPRETOTHIDOSPHADEVTACALLITEDLVDENKT
RESPECTATIONALLY STENKERDHILVOYD OWSSACM THIVERERGYMOCEE
VHILINETHOUSODEDINGREFELESPECCHOOT IN YNOPEREETRINGADEVACHEFTED
LORLIKKEDYDEVERGITVERAMETEDMOLYFDEPLAWGENFFELXDEKHPAAFVOEPSCD
OWLLIGGTEPINLDERNITVEVEFERMAGLLEKELOKVSGTFSAVFORKGLEISVWYNRESC
OWALKETTADDROTT

CPn_0490 571276 573333 CPNF hypotholical protein IMYNLLINAIDAADPORLABILERLEPHTYEEGEVETENTPAYFLGFHEISSE LOVNERG SLAQLAYEAVEAHHELINAARKEAREHYLEMEGED FAYAMEELLEPESFYCKLEAADDRRI. VIGHCYEARMFTHYDRYCDELEFSKKLEHFTYLEMEGED FAYAMEELLEPESFYCKLEAADDRRI. FLPLMSKSUTRPHLKIRKFLPL//GMVTDRPPVPEDHKI

POGLRHTANDILEPTTOESGDI/EFYGSTSEPIERIPI

POGLRHTANDILEPTTOESGDI/EFYGSTSEPIERIPI

EPYKEHSFFFYRDMLQE

PRISPSDERHAREIQKH

LEODPCPPFLKAMETOHITSOGVLFSRYFPSASLKOMFLSNYSRYYLOHIYFQIPSPTSG

EFFSNRDRSFLLDL/FAGISVFWADLESKRLLOY/IKRNNKDYGMFVPKHQAEQFAQSYFI

GIHGSCLIAGDYDEFLRELLTCMHTLSQOFTIPEFPPQTPLAILTGCGSGAMELANNVAT

ELSILSCGNLIGLDTTNAYVEAKMSYAIPDLLEROADFHVDLAVFVIGGMGTDFELLLEL

ISLKTCKKALVPVFLIGPVDYWKSKITALYNSNHAVGTIRGSEWVHNCLFCLSSAKAGIA

CPn_U491 574595 573336
CT389 hypothetical protein
ILSSLYTVFTMKTAFHSCYSWFCWLFSFLVLFVGGIAGGEPLCPDCKYETKSVLRSDQLP
ILSSLYTVFTMKTAFHSCYSWFCWLFSFLVLFVGGIAGGEPLCPDCKYETKSVLRSDQLP
HLWNYENDCYLTGYVQSLLDMHFLDSRTQVVIENRAYLFSLPVDSSLSEAITNFVRDL
PFICAVEICERPYGECITRSSAERPLLPKEKTLGMPIFCGKEGWWLPDNTILFSPLIADP
RQVTNSAGIRFNEKVVGNRVGATIFGGDFILLRLFDVSRFHVDCDFGIGGGYFSVFDLDM
PESGMVNSDFFVAGLWSGAIDKWSFRFRLWHLSSHLGDEFILTHPNFPRFNLSDEGVDLF
ISFRYTPQIRLYGGCGYIVSRDLTPPERPFYCEWGAELRPFGLREGNLHAQPIFAMHFRC
WEBQKFGLDOSYILGMEWAKFQEIGRKIRAVLEYHQGFSKEGQFIREPCNYYGFRLTYGF

CPn_0492 574643 574804
No robust homolog present in Genebank/EMBL as of 11/7/98
LFSLIFPICEERNSQQTYKHLHVESACFLLESPLKIHWSSPYGFPPFYRRDLKL

CPn_0493 575142 574855 No robust homolog present in Genebank/EMBL as of 11/7/98 SKTEGSHSKTSKGFVGRFVGWIKTFTGRGSKKRSPSSFSPTHPYIRLRTYTRSPKQSGVE RKQEDAETSFIETPKGILKKPGNKDPKGKHVHWKDS

CPn_0494 575370 575146
No robust homolog present in Genebank/EMBL as of 11/7/98
VINIRVNPYGSYRGRNPSFEDGKKDVPLSGNSRLHRRGGIRRKHKSASVGVTSGSKTGKA
SLEKKVKGISEAHFK

CPn_0495 575507 576793

aspc_aspattate Aminottansfetase

RRLKKNOKMAIQKAGAFLRCLPSESRPYLEHAMRNPHFSLLKPQYLFSEISKKLAQFRK
ENPEISVIDLSIGDTTOPLCRSITQAIKEFCVSQEKQETYRGYGPETGLEKLRYKIASEV
YERRISPEEIFISGAKPDIFHLFSFFGSERTLGLODEVYPAYRDIAHITGIRDIIPLAC
KKETGFIPELPNQOSLDILCLCYPNNPTGTVLTFOQLQALWYANQHGTVLIFDAAYSAF
VSDESLPKSIFEIPEAKYCAIEINSFSKSLGFTGMRLAMNVIPKELTYDNNEPMINDMKR
LFÄTTFMGASLLMQEAGYYGLDLFPTPPAISLYLTNAQKLKKSLETAGFSVHGGDHAPYL
WYELPEGISDEEAFDFFLHQYHIAVTPGHGFGSCGQGFVRFSALTQPQNIALACDRLCTA
SLKETMYLA

CPL_0496 576751 577812

CPL_0496 576751 577812

CTT591 hypothetical protein
PRIMFFKRNDGSCMTILRKLSGYLFFFSLFCSFIYVATCGSQPDSVSSPKIAIFLSFPH
PLIEDCSKSCIETLKDFENLPEIVVLNAEDSIVKARKIARSLHTDKNVVAIVTLGTIATK
WESLETGKPVIYAAVPDRESLTLFKNTMNIYGVNDTLDINQYCFAIQAVATNAQSIVYL
KPSEPFPSDLQKEIVKKLHASGIEVIEISITSSTFKTRIRQAIDKRPSAIFIPLSPLSHK
EGTSTLQEILKEKIPIITDDTSLISEGACIACSVDYKKSGKQIARIVHHLLYNNHDVDSL
RKITAQRLSPTTFFNEDIIKYLGIKLHKTERNQFLSFKSKKLEKSEKGKNVAVS

CPT_0497 578107 57820

CTT388 hypothetical protein

CPÄ_0497 578107 577820 CT388 hypothetical protein IFPRVVLDDSWILEVKVTPKAKENKIVGFDGQALKVRVTEPPEKGKANDAVISLLAKALS LPŘŘDVTLIAGETSRKKKFLLPNRVQDIIFSLHIDV

CPA_0498 579062 578085

No_EDDUST homolog present in Genebank/EMBL as of 11/7/98

YCRIRRAPFMNRKKARWVVALFAMTALISVGCCPWSQAKSRCSIDKYIPVVNRLLEVCGL
PEĀENVEDLIESSSAWVLTPEERFSGELVSICQVKDEHAFYNDLSLLIMTQAVPSYSATY
DCĀVVFGGPLPALRQRLDFLVREWQRGVEFKKIVFLCGERGRYOSIEEQEHFFDSTATY
PTEEDWESGNRVTPSSEEEIAKFVMVMLLIPRAWRDSTSGVRVTFLLAKPEENRVVANRK
OTLĪLFRSYQEAFPGRVLFVSSQPFIGLDACRVGQFFKGESYDLAGPGFAQGVLKYHWAP
RIGLETLAEWLKETNGCLNISEGCFG

CPn_0499 580404 579205

No robust homolog present in Genebank/EMBL as of 11/7/98
LSYVLLIFYECNCSTMSSVMQSSGTPNPEEVTSPESTEENKNVVSSDEAQATHAVALPIV
TQLSLPEGVGTSSEETASNPRVDEIVAEVSSSRAVADQISSLVERVGELLDDLKGAQSLF
TSFOSELKNCLPAWKSSTRALETRGAGDNADIARLELFRSDYEAVLGHANQFHGKAHLIL
SKLTDVHHKLOGLSREDLSLAFDNNDRVLEHLGSLGLDVDAEGKWSLSCERGIPRLVLTA
DSMLVOJKKVNLPTVEELBRTLCGTTESSSDPRVEESLSCCERLLNELERRIMNTVGFISS
CYDNIVFVLMWIVRRINLLPGLGCLPFHNPDASQEDQRSSSGERSTRRERLSRRSDLSEE
EMIVRAEGESIHPESPHGDGRNQPSRGDKQDSDSEEETEL

CPD_0500 580647 582362
pro5-Prolyl trna Synthetase
QHHMKMTSQLFYKTSKNANKSAAVLSNELLEKAGYLFKVSKGVYTYTPLLWRVVSYMMNI
IREELNAIGGQELLEPLLHNAELWOHTGRWEAFTSEGLLYTLKDREGKSHCLAPTHEEVI
CSFVAQWLSSKROLPHLAVQIATKFADEIRPREJEIRSRELLMEDSYTFSDSPDMNEQY
EKLRSAYSKIFDRLGLAYVIVTADGGKIGKGKSEEFQVLCSLGEDTICVSGSYGANIEAA
VSIPPQHAYDREFLPVEEVATPGITTIEALANFFSIPHKILKTLVVKLSYSNEEKFIAI
GMRGDRGONUTVVASKLNADDIALASDEEIERVLCTEKGFIGPLNCPIDFFADETTSPMT
NFVCAGNAKDKHYVNVNMDRDLLPPGYGOFLLAECDTCPENPGHPYRIYGGIEVAHIFN
LSTPYTDSFEVNFQDEHGOTQQCWMCTYGIGVGRTLAACVEGLADDRGIVWFMALAPFSI
TIAFNGJDTVSQELAETIYHELQSCGYEFLLDDRDERLGFKLKDSDLIGIFYKLILGKSY
QSDGIFEIEDRGGEKYTVSPEAFPTWCQNHLA

CPG_0501 582424 583650
Inca-4PFI Transcriptional Repressor
ILLTPGOSPP IMENTIVE/VOLEMARSKYSKRESKILDILFATTELYLKTG/PVGSKTLK
ESPGODIGITAT INNYFAELBALSFLKKNITTOGRI LETDLALRHYVDISØESCFEAELSAPI
FOK LOG/LIGEISKRI IKDIJKATELLSE ILDJEPFESOPPFENDSVTNÍ QUTYOPKGRATLLTGEFOG FETDTLMLIGEACOTLISTIKR LEKFLANY FRED TREEDIKKEEHLSMSLYNEV
VVRITTRYCHESSEDLYOTEMIKLLKYEAFKOPEVLAUDLISLEFENIRGMSELLHIGMIKG
RATAF LÖKEIJB I HATGERING SCHITT DYYMIRSPLIALS ILDJE INLYKEALPILKLEAN
KIMETLINGSPYKEKLSFRRPLITINGKLISNEP LLTTEY/CSTKLLPYKETI.

CPn_0502 GPDE-HSP-70 COLDCTOT GDVMTDTPPENEOHESN/ONENE/EHLDOETVTLKTELKEKNDKYLMALAESENSRKRL GKERQELMOYALENTLIDFLNPIESMEVALGFATOMSDDVKNWALGFMMILNQFKQIFEE KGIIEVSSIGOKFNPFLHEAVQTEETSEVPEGTILEEFAKGYKIGERPIRVAKVKVAKAP TOPENNE

CPn_0504 586418 588514

vac8-ribonuclease family

ATOPTSETTGFLVCCFKLTGGAQLLKKPKRKPGRRTYGKSLKIFIPGTLFVHARKGFGFV

SPDNPEEYPFDIFVPARDLRGALDGDHVIVSVLPYPRGGQKLKGTISEVLARGKTTLVGT

ITSLVSFTSALAYTSMSSGSLIPVELLPGRTYKIGDRILLSTPPWDDKPGEASPALOM

LEFIGHITHNAKDFFBAIQAEYMLAEEFPPEVIEEASLFSQKHITOVLHSRKDLRDLLCFT

IDSSTARDFDDAISLTYDHNNNYILGVHIADVSHYVTPHSKLDKEAAKRCNSTYFFGKVI

PMLPSALSDNLCSLKPNVDRLAVSVFMTFTKSGHLSDVOIFRSVIRSKYRMTYDEVNII

EKKHSHPLSKILDEMATLSKKFSDIREERGCIRFVLPSVTMSLDNLQEPVALIENHGTFS

KKLIEEFMLKAYEVVAYHISHGGVSLPFRSHEPPNDENLLAFQELAKNMFDITFTPTQE

PDVGYLLOTTSAGHPLEGVYLNSOFVRSMKTASYSTENKGHYGLKLDYYTHFTSFIRRYID

LIVHRLLFNPLSIDOTHLEIIVRACSTKERVSAKAENSFENLKKTRINKFLOROPKTTY

HAYIITANHESLSFVVTEFCHEGFIAAAELPKEYSLKKNALPESIPDKMKPGASIKVTID

SVNLLTQKIJWSIATTTEDKPKKKKKTFKKKTKKRS

CPL_506 589055 589840
CT_371 hypothetical protein
CPM_EISPIPRREGKSFILMILKLYSKETNAHFLISCRRIMKKYFITGLVILLPLAITIAI
CPM_EISPIPRREGKSFILMILKLYSKETNAHFLISCRRIMKKYFITGLVILLPLAITIAI
CPM_EISPIPRREGKSFILMILKTYLLGFILTRIMI
FK_ELLSIYDKILHRIPIIKTVYKAAQQWMTIIFGSKSGSFKOVVMVFPRNANQCIGLVA
KJAPTVCCTGEKEDDPLVTVFIPTPNPTSGFLTLFRKSDIVFLDMKIEDAFKYIISCGV
LSTPMACPSSPLPDELHQDQGS

An_0567 589898 590122 CM21.1 hypothetical protein STYPPOFPLSGEIKKFNIELFMTRMSKQARRRAKSPKKRKPKYAIVHPAPAPRIVYKLHT NALSTSDSIFIPKIG

CPn_0508 590133 590300 CT421.2 hypothetical protein SRIMSRHRSYGKSVKGVTKRNVLKRFERVEVLRKLGRWNDSTAKKVTGLPKTPILK

CPn_0509 590299 590808
(predicted Metalloenzyme)
NKFVFLYGNFIRVTQEKIKIHVSNEQTCIPIHLVSVEKLVLTLLEHLKVTTNEIFIYFLE
DKALAELHDKVFADPSLTDTITLPIDAFGDPAYPHVLGEAFISPQAALRFLENTSPNQED
IYEEISRYLVHSILHMLGYDDTSSEEKRKMRVKENQILCMLRKKHALLTA

CPn_0510 590804 591973
tlyc-cbs Domains (Hemolysin homolog)
QLNMLHILLAIFCILLFLAFGLTOPSCHGSSKFLKTLNQRFFKDKGREYPPFPSAPTILA
TLLCILYGALGTKLYTLLPPKTAHKDLLFWPLYSLSALIAYGFLPPWISTKVPKETTAHL
RFLASVPOLGLFFLQLLFYRRPNQQVRSSTSFQSQLSEALSAFDNLIVREVMIFKVDIF
ALPEETIQEALVLVSEEGYSRVPVYKKNLDNITSILLVKDLLLLYTSSHDLSQPISSVA
KPPFYAPEIKKASSLLQEFROKHRHLAIIVNEYGFTEGIATMEDIIEEIIGEIADEHDVQ
ENTPYKKIGSSWIVDGRMIISDAEEYFRILKIDHENSYDTLGGHVFHKVGAVPQKGMRIHH
ENFDIEIITCTERNVGKLKITPRKRKFHIS

CPn_0511 592141 592488
rsbV-figma Regulatory Factor
MSDIOKEEHGSTTIFHLHOKLDGISSPEVQENISQSLAAGSKNIILDCAHLDYMSSAGIR
VLLOSYHQVQOHSGKIVLTTVPKTIEQTLYVTGFLSYFKIFNTVDEAIQTLNKDGD

CPn_0512 592538. 594412
CT425 hypothetical protein
SLPLTMRRSVCYNPSIARAGGISTWKFLYSLATPLPAGTKCKFDLAGSGKPTDWEAPAT
SLPLTMRRSVCYVNPSIARAGGISTWKFLYSLATPLPAGTKCKFDLAGSGKPTDWEAPAT
DLSOTRIVIYAEMPEGE I EATA I PVKDHPVPOF EFTLPYELQVGETLT I VMGASPNHPQ
VDDAGNGAQLFAQRRKPFYLY I DPTGEGNYDEPDVFSMD I FGNVLKK LEIFTPSYVVRNK
RFDITVRFEDEFONLTNPSPEETR I ELGZEHLRENLIMQLFI PETGFVI LPNLYFNERGI
RFI CLKNELTOG IFT I SAPT I KTGFADSAFILMMGLLEGESERVDSEEN I ETCRRYFRDRAL
NFYAGOGFENGENLITD INKL I NOTVOLFNEEDF FITLUSFOYSGEPHLEGVRHI LHTKE
TKSIGZHKEZKH I LAKLYKSTVNHDM IS I PISFTASKEHGFDFENFYBEFENVEL YNAW
SKSICTTALLINNPFP I COKKNEOPRITV I EGLKKI LLRFOFVAGGLDRG I YKDYFDS PQVQ
YSKGLTAL I CHKYTRESLAFALLFARHEGYATTUFF I VLSFNITGAPMGSELSTGSK PGLNV
NRIHZSHVATTALLKTVE I I RNGEVLHTFFFDCHHILDYEYDDMVPLSSVTLKDPNGKAPF
VFYYLRVTGADNAMARSSTLWGLIN

CPH_0513 944515 95475;
Fe 3: Oxider shert are very statement of the control of th

TK [KEYDIO] III [KALTA] EYAYLISDLUNLS IRUVLLT LOS I POGGAE ILVDK IRN FLAPKRICHS SEPILNI HKMANIQLG (HSNITMLEYHKEO HMWKYRDLQDETQGFKN FILLKFAQENNYLCKRLRKGGCHA I PLKSLMAVAR I P. SMKALMYLG I EAALDL LSCGANDLGGTHMGEKYFQMASSKEP I KMDAEGMAALI TOQGRTPCLTNSSHV

CPn_0514 595690 596520
CT427 hypothetical procein
CMCOPHTTREMAMINULOPISLOCOSY/INSFPLSLQLIKRNDIRCVLAPPADLLILLI
CMCOPHTTREMAMINULOPISLOCOSY/INSFPLSLQLIKRNDIRCVLAPPADLLILLI
CMCOPHTTREMAMINULOPISLOCOSY/INSFPLSCODILLICANA/TEPEDIS-PEDIA/TLEGE
CMCOLTALPHIA/METHORITY/INSTPLSCODILLICANA/CHICH/CPT-TYDL
CMCOMPULTRLPFVFALLICHTS/MERHPLENLAMERALQOFESSPEEVLKEAHQHTGLPPS
LLOEYYALCOYRLGEEHYESFEKFREYYGTLYQQARL

CPn_0515

CPn_0515 596450 597181
ubiE-Ubiquinone Methyltransferase
EKNTTKALKNSONIMEPSTNKPDCKKIFDSIASKYDRTNTILSLGMHHFWNRSLIQILGS
GYSLLDLCAGTGKVAKRYIAAHPQASVTLVDFSSAMLDIAKQHLPGSSCSFIHSDINQLP
LENHSYPLAAMAYGLRNLSDPHKALQEISRVLMPSGKLGILELTPPKKTHPTYSAHKLYL
RAVVPMIGKSVSKDPDAYSYLSKSIQQLPKDHDLEDLFSKSGFYIAKKKKLFLGAATIWL
LFKO

CPn_0516 598904 597255
NO robust homolog present in Genebank/EMBL as of 11/7/98
NISISFRVSWFVKIILAVLGRAIAKAYYYCMYARGLCDFPTLVPNERLPIGPFFVPOHTS
GAKGKEFAKRNFSIISGLDDILKLCILQRRPFALQWONLSVKSDYEEAGPAIGIRSLEPQ
VSQISPAHGRLCSTLVQWAPPILGSEEQLVWLEETHKRIKFPKSIGSKDAVIVDSEMVFVN
ANPYOEIPAASETVESSPVARGNITDTMPAASGTTDTTSGVSEAAAAEAAVDSTRGTEEE
PSFSLRYALVVQNVPYPEPPKEPFVMFTDEEKSLILEATRARRMELDLYNGYLADYELSK
DEIOKHVPDLPENMRTNMRWSERLYKFFFKTKKEGLEBIFINKELGNMILARGLAATOSQ
ARIKVTNSLVAMLLOSFNVGRSCTAKPLPTSKLDLFKSEFESKPKNNILTEFLVASDEEI
LFKGLRVLEFGIEGWYDHPDQAGEIRSVLEGLVQAGRISGYWENQPFGRFVLRGVGERT
ELVELLESLVASGEIMQFFESSDEEGAFIIDNEPSKTAMLKQRFKSCVRTKLVGSFADES
LPRGRFTILV

CPn_0517 599637 598795

No robust homolog present in Genebank/EMBL as of 11/7/98
FIMSSLLSCGRIEPTRVTCSLKTYLEDTSQNQLSTRLVRASVIFLCALLIILVCVALSSL
IPSIMALATSFTVMGLIIFVMSLLGDVALISVLTYSTVTSYRONKRAFEIHKPARSVYYE
GVRHWDLGRSSLGTGEIPIVRTLFSPFQNHGLINHALAAKIFLFMEHFSPEPPNEPLVDWA
CLIRDFRPHVSSLCTVIEKQGSSLRYKEGNTICEAFRSDYDAHFAMVDCYRLIHSKLIIE
KMGLKNIDIIPSVMVREDYPSRPGEGYREGLLRMYGKGKAL

CPH_0518 600806 599832
CPH_0518 600806 599832
CPH_0518 600806 599832
CPH_0518 FOR C

CPRE.0519 601707 600904
dåpf-Diaminopimelate Epimerase
OPFKLRILVYWMAFYSPSTISKYFIVSGAGNRFLLGETLPEVEDVRFLCQETR/DEFLYLV
KPŠSCADAQLIIFNSDGSRPTMCGNGLRCAIAHLASOKGKSDISVSTDSGLY\$GYFISSM
RYLVDMTLADWRASVHRLESRPDPLEREVVCIHTGVPHAVVILPEISTDLDIALDFLRY
HÖTTSPGDYNVHYVOILGHCQLRVRTYERGVEGETAACGTGALASALVVSNSYGWKESIQ
IBMCGELMTVSQNRGRVYLQGSVTRDL

CPR-0520 602233 601646
clop-clp Protease
gamyfmadgeviklabilekellearryffsepyteksasdaikklwylelkdpgkpiyf
gamyfmadgeviklabilekellearryffsepyteksasdaikklwylelkdpgkpiyf
vinspegsydagfarwddikhlispyttyvtglaasmgsvlslcaafgrafatphsrimi
h@FSIGGPITGQATDLDIHAREILKTKARIIDVYVEATNQPRDIIEKAIDRDMWMTANEA
kDEGLLDGILFSFNDL

CPT-0521 603803 602241
gipa-Serine Hydroxymethyltransferase
KSLKVPEKFKKFAIVEIFTKVVAVVSLHKK!LENASKKKOGLASTAYLAALDHLLNAF
PSIGERI IDELKSORSHLKMIASENYSSLSVQLAMGNLLTDWYCEGSPFKRFYSCCENVD
AIEMECVETAKELFAADCACVQPHSGADANLLAVMAILTHKYOGPAVSKLGYKTVNELTE
EEYTLLKAEMSSCVCLGPSLNSGGHLTHGNVRKVMSKLMJCFPYDVNPDTEECFOYAEIS
RLAKEYKFWYLIAGYSSYSRRLNFAVLKQIAEDCGSVLWYDMAHFAGLVAGGVFVDENP
IPYADIVTTTTHKTLRGPRGGLVLATREVESTLNKACPLAWGGPLEHVIAAKTVALKEAL
SVDFKKYAHQVVNNARRLAEFFLSHGLRLLTGGTDNHMYVIDGSLGISGKIAEDILSSV
GIAVNRNSLPSDAIGKWDTSGIRLGTPALTTLGMGIDEMEEVADIIVKVLRNIRLSCHVE
GSSKKNKGELPEAIAQEARDRVRNLLLRFPLYPEIDLEALV

CPn_0522 603825 604655
CT433 hypothetical protein
REPLSPEKTSLAFKVKNVNQRMIKKNOGKKKNYFØYIPLKVQKLRQPSFYPKRLMTLYLG
LNOKTARKYQAHYLPILTLFFYAKSTPONKRALØFLPJATHVILTSPSSTHLFLSRMTSL
LSYATLKTRYLCIGESTKERLLSFIGGVKYVYÄRGEIABGIFFLLQALPSSARILYPHS
SLARPVIRSFLYNRFIFFSYPHYTVKPRKLKKNILSKYKKIIFTSPSTVRAFAKIFPRFP
EKTYWCQGRMTLQEFQKFSSQKQVSLLETLGKSRTSP

CPn_0523 604720 605052 No robust homolog present in genebank/EMBL as of 11/7/98 PMAGSAPFOFDOTAPSIGEPPATRERYNEFFLALFVTIAIALWHALIATTIAIGLCIHPLC SFIFUTAIFLYFISRYICSHYARNVYIALDVVFDHOKLODMRSHSPIFGDR

CPR_0524 605079 606179

RE CODUCT HOMODIST DEGENET, IN GENEVALEMBL AS OF 1177708
PYFYERIAM CONTRACTOR DEGENET, IN GENEVALEMBL AS OF 1177708
PYFYERIAM CONTRACTOR DEGENETATION OF TAKE OTTALLAGIALALIVA
LOCATOR GENEVALAL OTTALE AND TELEVIRONTE EVICAN WORKN
VICTORING EYNORMEN PER VA OTTALE OTTALE AND THE VICTOR OF THE TELEVIRONTE LEVEL OF THE VICTOR OF TH

CFn_0525
CT198 hypothet ical/s
GIIFMIDALLSILAJOEDIT/MIRLMRVKKEHQKELAKVQSLKSDIRRKVQEKELEMENL
GIIFMIDALLSILAJOEDIT/MIRLMRVKKEHQKELAKVQSLKSDIRRKVQEKELEMENL
KTQIRGENRIQEISEGINKLENQQAAVKKHDEFNALTQEHTTANKERRSLEHQLSDLMD
KQAGGEDLIVSLKESIASTENSSSVIEKEIFESIKKINEECKALLEQRTELKHATNPELL
KQAGGEDLIVSLKESIASTENSSSVIEKEIFESIKKINECKALLEQRTELKHATNPELL
SIVERLLANKKDRVYPIENRVCSGCHIVLTPQHENLVRKKDRLIFCEHCSRILYWQESQ
VNAQENSTAKRRRRAAV

THE MADE AND ADDRESS OF THE STATE OF THE STA

CPM_05¢7 609910 608726
sucB-Dihydrolipoamide Succinyltransferase
sucB-Dihydrolipoamide Succinyltransferase
surHyferrpkigetssggsivæmiknigdhvardeplifvstdkiatelpspkagrlvr
fcvybdbevasgdviglieleeiseaddestsceptscetkseagssssswfspavi.sl
agrafiglonigkiagrakggrvtrodleavisesogvsipeifogevnipmsplrrai
asstlskssdevphaslvvddvvtdianlisserorfluthgvklitisfivollagtir
fyllnsslogttivkksvivgvavniksegvvvpvthenodbeglvsiakaladlssrar
likkldpsevgcsvivtnfohtgaligmpirrypevailgigtiokrvvvrdddslark
hvyvtltfdhrvldgiygsefltslknrlesvtmg

CPn_0528 611165 609921

gltT-Glutamate Symport

LMKLMMKIFIGLFVQVTLGLVLEDKAIFFKPIGDIFLNLLSMVVYPLVFCSMVLGIASIS

LMKLMKIFIGLFVQVTLGLVLEDKAIFFKPIGDIFLNLLSMVVYPLVFCSMVLGIASIS

DMKKLGRIGIKSVGLVLGTTALAIVIGLCFAWIFSPGNCCDFAQAQSMDSAVTVIDSNKT

ANYTLSIIAQVFPSNPVRSFAEGNILQIIIFAIFLGIALRLSGERGRPVERFIDDFSEIM

LRMVNNIMSFAPYGGASMAWISGNNGLGVLWQLGKFIIAYYLACLFHATLVFGGLVAFG

CKMSFSKFLSSMDAISCAVSTASSSATLFVTMRCVSKNLGVSAEVSGFVLPLGATVNMN

GTAIFGGMAAVFIAQAYNCPLSLSSLLLLUVTATFSAVGSAGVFGGGMITLGSVLASVGL

PIQGIAILAGIDRLRDIVGTPMNILGDAVVATYVASGEGELSPYESIKQESVETT

CPn_0529 612298 611165
ycah-ATPase
FSCKEIRAFKRGTMKKRFPSTLFLFYRRVTIAISLEGILGMGMLGSLLSKVFAFLVACMN
FSGKEIRAFKRGTMKKRFPSTLFLFYRRVTIAISLEGILGMGMLGSLLSKVFAFLVACMN
KESMSTPYRARSTVISVGNTVVGGAGKTPTVLMLAEALRLRGYSCGVLSRGYKSQSSRQK
KLTVVDSKVHSASYVGDEPLLMAEKLPEGSVWHKDRRISAARAAEKFGILLLDGGUYK
KLHKDVEIAVVMCODPLGGRAFFPKGRLRDFPLRLKTVDAIIVMCGKEAGTVVKRVSNA
PQIFVKETIASVWTHNGERIPKEALRELRVGVFCGLGFPCGFLAMLREEGIHILGKYLL
PDHAAITKKELNYFCQMAMRQGQGLLCTEKDSVKLPRLSGEVSLLPIAKVEMRLSVNQD
DTLSLLNMTEQIHKNRGN

CPh_0530 613323 612460
spoU-rnA Methylase
svvlmgkflmrRcGslafwefcsmdCigkhnplvkealalkrsrcrksswflvegareig
kalrtgyllqhvfcsthlsekekeflytelkrnstkilycldstlaglsfkehdbfyvd
kavmnkepfligrrnagpfylliegvekrgnvgailriadgasyddyllchpiydlyn
nvrssigavfslpilsseedkelfkgesmtytvtspraftmyfsknylgptalvfgs
ekdgltedmfsedfseialpmlgesdslnlatsvaavayevvrgrwyn

CPn_0531 614198 613245
SAM dependent methyltransferase
DSSKDDFRKEKGRRKSQYRDRYVNKDTGRHSKTYFSLIRERLVMDYKLLDSGDGNKLECF
DSSKDDFRKEKGRRKSQYRDRYVNKDTGRHSKTYFSLIRERLVMDYKLLDSGDGNKLECF
GPVTLIRESSIAVWPKSRPELMSQAQLQYVREGERGAMKNFKRLPEBWEVAFSDVRCLIK
RTPFCHLGVFFEHNGFWPALKQAIEKHKERQVLNLFAYTGAGSIFAAKCGARVTHVDASQ
AAVRWAQRNVEKNAFPERIFFWIEDVISFLKKEIRRNKKYQVILLDFPSYGRGPDGEVF
KIDKDLFPLLSLCSKLLADDASYFLLTSHTPGHTPEFLRAIARRSVPTLVSEAWSCGESF
CGEGVGALPSGSFVQWIA

CPn_0532 614716 614075
ribC/risA-Riboflavin Synthase
eSFCCKDSVVKWGGMFSGIIQELGEVCFFEAQGNGLSLGIKSTPLFVTPLVTGDSVAVDG
VCLTLTSCNESKIFFDVIPETLACTTLGEKRCSDQVNLEAALKWGDSIGGHLLSGHVFGT
AEIFLIKENRYYFRGSKELSQYLFEKGFIAIDGISLTLVSVDSDTFSVGLIPETLQRTTL
GKKREGERVNIEIDMSTKIQVDTVKRILASSGKD

CPn_0533 614918 615385
CT406 hypothetical protein
EVAPMCPFCNHGELKVIDSRNAPEANAIKRRRECLKCSORFTTFETVELTLQVLKRDGR
YEHFQESKLIHGLNAASSHTRIGQDQVHAIASNVKSELLGKQNREISTKEIGELVMKYLK
KADMIAYIRFACVYRRFKDVGELMEVLLSATPDMEK

CPn_0534 615389 615784
dksa-Dnak Suppressor
LNFTRSKVPLSDDEIEOFKKRLLEMKAKLSHTLEGNAQEVKKPNEATGYSQHQADQGTD
TFDRTISLEVTTKEYELLRQINRALEKINESSYGICDVSGEEIPLARLIAIPYATMTVKA
QEQFEKCLLSGN

CPn_0535 615763 616296
LspA-Lipoprotein Signal Peptidase
KRTPIMKLSSMATRERSTLLVITLEVILDMYTKLVVLLQYKDLQILTHPTLYTHSWGRFS
FSIAPVERGAAFGLESNKYFIFLERIFVILGLLAYLFFKKKSIOSTTQTALVLLCAGA
IGNVGDIIFYCHIVDFISFNYKÇWAFPTFNVADVLISLGTLLLVYKFYFPTKQTEKKR

CPn 0537 617815 618189

CTRIALL hypothetical protoso LIFELEMONYLLOGLIFCCVLLDIGMCTIFVMTICFLE LAPLMLNKKLLCGWLKKKKNRGGLSEDIDELLDEKKOR

LDOG I KWCAALVL IWKV

618511 618128

CTB14 hypothetical protein TREINFACHWOSPOKCFVIKIKTERDIYMFRNNHKPKKTKCKRFRWLRGVLFGGFIATLL TREINFACHWOSPOKCFVIKIKTERDIYMFRNNHKPKKTKCKRFRWLRGVLFGGFIATLL TREINFACHWERTIETKER GETTER FYRREIL POTTROFFYAKIGARHIGHELDEK

CPn_0539 618678 621545

pmp_19-polymorphic membrane protein
GYNLIGLRHMKOMRLWGFLFLSSFCQVSYLRANDVLLPLSGIHSGEDLELFTLRSSSPTK
TTYSLRKDFIVCDFAGNSIHKPGAAFLNLKGDLFFINSTPLAALTFKNIHLGARGAGLFS
ESNVTFKGLHSLVLENNESWGGVLTTSGDLSFINNTSVLCQNNISYGPGGALLLGGRKSK
ALFFRDNRGTILFLKNKAVNQDESHRCYGGAVSSISPGSPITFADNGEILFGPENBEELGG
ALYNDOGAITFENNFOTTSFFSNKASFGGAVYSRYCNLYSQWGDTLFTKNAAAKVGGAIH
ADYVHIRDCKGSIVFEENSATAGGAIAVNAVCDINAQGPVRFINNSALGLNGGAIYMQAT
GSILRLHANQGDIEFCGNKVRSGPHSHINSTSHFTNNATTIQGAPREFSLSANEGHRICF
YDPIISATENYNSLYINHQRLLEAGGAVIFSGARLSPEHKKENKNKTSIINQFVRLCSGV
LSIEGGAILAVRSFYQEGGLLALGFGSKLTTQGKNSEKDKIVITNLGFNLEDLDSSDPAE
IRATEKASIEISGVPRVYGHTESFYENNEYASKPYTTSIILSAKKLVTAPSRPEKDIQNL
IIAESEYMGYGYQGSWEFSWSPNDTKEKKTIIASWTPTGEFSLDFKRRGSFIFTTLWSTT
SGLNIASNIVNNYLNNSEVIPLQHLCVFGGPVQIMEQNPKQSSNNLLVOHAGNNYGR
VMKHVFPXKGTSRGSWRNYGWSGSVCMSYAYPKGIRYLKMTFPVLQYTKLVQNFPVETG
YDPRYFSSSEMTNLSLPIGIALEMRFIGSRSSLTQVSTSYIKDLRRVMPQSSASLLVINN
YDPRYFSSSEMTNLSLPIGIALEMRFIGSRSSLTQVSTSYIKDLRRVMPQSSASLVINN
YTWDIQGVPLGKEALNITLNSTIKYKIVTAYMGISSTQREGSNLSANAHAGLSLSF

CPn_0540 621631 626862
pm_20-polymorphic membrane protein
fillylsliefynisdrfssmkwlpatayfaavlbaltafgdpasveistshtggdpt
fillylsliefynisdrfssmkwlpatayfaavlbaltafgdpasveistshtggdpt
sdaaltgffosstetdgtyytivgditfsffinipvpvvytpdandsssnsskggssssga
savaadprogaafysneantilfftdsongslildnikmigddalyskgplyftgl
knitftgnesoksggaaytegalttgisderofslindalwingddalyskgplyftgl
knitftgnesoksggaaytegalttginickiefisdroggatyvkgatlfnalds
kfeknisogagggityestliisnitksiefisnkasvpapapeppspapslinsttid
kteintsogagggitytestliisnitksiefisnkasvpapapeppspapslinsttid
tstlotraasatpavapvaavtpteistgetakoggityakggisistfkdltfksnsas
vdatutvosstigesgaittakaltinsgallttesgrtanksgggityavgovtledia
nieminnickgeggaittkaltinngallttesgrtanksgggityavgovtledia
nieminnickgeggaitkaltinngallttesgrtanksgggityavgovtlidli
tskelfoenssekhogggavvtktfocsdshrloftsnkaadegglycgddvtlthi
tokelfoenssekhogglslasgksltmfilesfclanktakenogganvpenivlitti
tokelfoenssekhogglslasgksltmfilesfclanktakenogganvpenivlitti
tokadktocsftyitnvnitnntatgnoggiagkahdroantinvosnoakkoggvul
edaellekvitgsvsontatesggityandlolalesftitdnkvesltstinlog
glinsgavtltnisgtfgitgnsvintatsodaloggiyattslsincompilifsnn
saatkktstikolaggaifsaavtiennsopiiflinnsakseattaataankdscogaia
naestltnipettfkonyaetggalocioltngsprrksiadnosvlfodnsalnrgga
ivgetidisrtgatfignskhogsaiccstaltlaphsglifennkvettattkasin
nlæalygnnetsdvislsaenosiffknnctatnkycsiadnosvlfodnsalnrgga
ivgetidisrtospotiilsaenosiffknnctatnkycsiadnosvlfodnsalnrgga
ivgetidisrtospotiilsaenosiffknnctatnkycsiadnosvlfodnsalnrgga
ivgetidisrtospotiiltnippesskerosponskokoplitennkvettattkasin
vlæalygnnetsdvislsaenosiffknnctatnkycsponstythonatvapptikl
systnadskokiditotyvlldprognlynnylgedritenidnsportpronatvapptikl
systnadskokiditotyvlldprognlynnylgedritenidnssgavtatnvvvlo
gelgakkoylctwnldprssoksiskkesgskolsponstythgrytnavdapperge
filgaafsovfehaeseyhldnykkkgsghstoaslyagnifyfpalrshyldfranddarpge
filgaafsovfehaeseyhldnykkkgsghstoaslyagnifyfpalrshyldrildnand
verlitentyfildensomstawldensomstavlaptildnandellingkfyflodnandellinnarfedyblengensomstavlaptillinnandellingkfyflodnandellinnarfedy 626862 621631

627137 628003 CPH_D541 627137 628003
SOLUTE binding protein (-yebL-Synechocystis Adhesin Homolog)
NNRSSYCTAFVMHKVIVFIFLTLYSLKSYGNDVIDKPHVLVSIAPYKFLVEQIABETCFV
YA LITNHYDEHTVELPPQQIKELRQGDLWFRIGEAFEKTCERNLTCQQVDLSQWSLIQG
KPCCNQHTTNYDTHTWLSPKNLKVQVETIVTTLSKKYPQHATLYQSNGEKLLLJLDQLNE
EILTITSKAKQRHILUSHGAFGYFCRDYNFSOHTIEKSSHVEPSPKDVARVFEDIEQYKI
SSVILLEYSGRRSSAMLADRFHMHTVNLDPYAENVLVNLKTIATTFSSL

628737 628000 CPN_0542 628000 628737
ABC Transporter ATPASE
FMTIRILAEGLAFRYGSKGPNI IHDVSFSVYDGDFIGI IGPNGGKSTLTMLILGLLTPT
FGSLKTFFSHSAGKQTHSMIGWYDGHFSYDFCFFISVKDVVLSGRLSQUSWHGKYKKKDF
EAVDHALDLVGLSDHHHHCFAHLSGGQIGRVLLARALASYPEILI LIDEFTINIDPDNOOR
ILSILKKLNRTCTILMVTHDLHHTTNYFNKVFYMNKTLTSLADTSTLTDQFCCHPYKNQE

628710 629603 CPn 0543 CPR_0543
(Metal Transport Protein)
KSGIFMLSSLIRDSFPLLILLPTFLAALGASVACGVMCTYIVVKRIVSISGSISHAILGG
IGLTLMIQYKLHLSFFPMYGAIVGAIFLALGIGKIHLKYQBREDSLIAMIMSVCMAIGII
FISRLETFROELINHLFGNILMVTBSDLYSLGIFDLLVLGIVYLCHTRFLALCFDERYTA
LNHCSVOLMYFLLLVLTAITIVMLIYVMCTILMLSMLVLPVWTACRESYKMTRIMFISVL LN1LCSFSGIC1AYCLDFPVGPTISLLMGLGYTASLCVKKR/NPSTPSPVSPEINTNV

CPR_0544 630598 629525

yhbZ-GTP binding protein
RESYPTCH INDICATE PROTECTION
RESYPTCH INDICATE PROTECTION
RESYPTCH INDICATE PROTECTION REPLANDED GO GO ATTINITION OF THE PROTECTION REPLANDED GO GO ATTINITION OF THE PROTECTION REPLANDED GO GO ATTINITION OF THE PROTECTION OF THE PROT 630598 THE HALLOW PERMANENCH PRINTER THE METALLE PROPERTY AND ARTER AND A KOMOVALNK ODDIGEDEGEE GOSFOKREPSYTEV LOGINGEOVICILYRFETORIAV

CTH_0545 030908 030633V 1127 CEV cibodomat protein TREAMBRIVANAIMMOS ADBRINDOMOSMOSMOSMOSMOTT CHLOMORGERWIIPAONOG ROBERTLEALANG FUNKKERIETY CONFEQU

430912 CPn 0546 CPR_094M 121-L21 Ribosomal F LEKORLTLSIERFIRKKLMEPYAVIOTGSKQYQVREGODVIDVELLGE/AGDKEVIFQDVL FVFDGTKASLGSPTIANAQVKAEYLSHVKGEKVVAYKYKKRKNYHRKHGHRQKYLRVKIR

26

_0547 5315R 632188 yobB family ALMORDNEZOLDE DEWEMENTE EUR JUHR ELFEGSTERN FRA ENALDWITHE ELF KVADECU, TEGUTDOGIYLE SRPKFLCKLSALRONTAQVMNUTFTDIGITATSGEGLSD EALKSLKPNGKISHVAITIE FGCGDGVQCFC'/LTVMEYC

33234 632191 CPH_0548 631234 632191
cysJ-Sulfite Reductase
kmylokfkaqovplvtrellscsdsindsdpiyrmvfdsndttisykvgdalgvlpens
kevsehvloligyspttlvnvkkrsekvsaqkfiggyvdldkipaklnsffpdkdpkitl
ydalgevrpoipielyaesvfpllprfysiasspdlhpksiellvkhvsypgkyqkkfgv
cssflcselqvndsytifvoptkhftlstotegkplumigactgiapykafleerlfnkd
pgnnllffgerkekvnfyyrefwnhaeeegklklflafsrerdokyvvqdllriqkdevr
kayeeggfffvcgkvlgievkhaleeilgkdtlaslrkehryvvdvy

CPn_0549 633662 633255
rs10-s10 Rikosomal Protein
PODVOHOPHNOHSLLRFLKKFKKRLLRSKGCMKQQKQKIRIRLKGFDQGQLDRSTADIVE
TARRTGARTVYDFIPLPTKREVYTVLRSPHVDKKSREQFEIRTHKRLVDILDPTGKTIDAL
KMLALPAGVBIKIKAA 633255

CPn_0550 635688 633580

fusa-Elpngation Factor G
LNYGENNKYMSNQEFDLSAIRNIGIMAHIDAGKTTTTERILFYAGRTHKIGEVHEGGATM
DMMAQEDERGITITSAATTVFWLGAKINIIDTPGHVDFTIEVERSLRVLDGAVAVFDAVS
GVEROSETVWRQADKYGVPRIAFVNKMDRMGADYFAAVESMKEKLGANAFPVHCPIGSSE
GVVGHVDLISCKALYFLDDTLGAKWEEKEISEDLKERCAELFANILEELATIDESNEAFM
MKVLEDPDSITEDEIHQVMRKGVIENKINFVLCGTAFKNKGVQOLLNVIVKMLPSPLDRG
NIRGINLKTDDEISLEPRRDGPLAALAFKINTDPVVGRITFIRIYSGTLKKGSAILNSTK
DKEENISKLLEMHANERTDRDEFTYDGIGACVGLKFSVTGDTLCDDNQEIVLERIEFFDP
VJØMAIEPKSKGDREKLAQALSSLSEEDPTRVSTNEETGOTIISGMGELHLDILADRMI
REFKVEANWGKPQVSYKETITVSGNSETKYVKQSGGGGGAYHVCLEIEPREPGKGMEVVS
KIVGGVIFKEYIPAVIKGIEBGLANGVLAGGGLVDVKVSIVFGSYHEVDSSEMAFKICGS
MAVKDACRKAKPVILEPIMKVAVITFEDHLGDVIGDLNRRGKILGQESSRGMAQVNAEV
PLSEMFGYTTSLRSLTSGRATSTMEPAFFAKVPQKIQEEIVKK 635688 CPn 0550

636174 635698 rs7-s7 Ribosomal Protein
MYMSRRHSAEKRDIPDDPIYGSVILEKFINKVMMHGKKSVARKIVYSALERFGKKLNLEN
VLEGFGEALENAKPILEVRSRRVGGATYQVPVEVASERRNCLAMQWIIKHARSKPGKSME VGLATELIDCFNKQGATIKKREDTHRMAEANKAFAHYKW

636698 636219 GPH_0552 6J6698 6J6219

£\$12-512 Ribosomal Protein
IQAGYVPSSSENKPLETKRALLYISMLVVVRLKREEYMPTINQLIRKRRKSSLARKKSPA
LQKCPOKRGYCLQVKTKTPKKPNSALRKVAWVRLSNGQEVIAYIGGEGHNLQEHSIVLIQ
GGRVKDLPGVRYHIVRGTLDCAAVKNRKQSRSRYGAKRPK

CPn_0553 637753 636812
No robust homolog present in Genebank/EMBL as of 11/7/98
GCMWRVVLRFLIIFILGRAVFPLRASESFSWETSTCLTVLGIPFIDIILTTNEDFVAQCG
LOIGTISSTNNAKIKEIFLIYKEKFPEASISFKRKEPLNLSQSHLSDLGILCHRNGETYA
EGMANKENGPALKOPKDLRLVLRCPNOPDTLLYSSEABKGIETNTCLENOGYTLLDGQL
ILYGDSIEKFLKETKRKNHTLVDLCDSQVVTTFLGFRYSLLNYVQVLFLSEDSAKILAG
IPDLAQATQLLSHTVPLLFIYTNDSIHIIEQGKESSFTYNQDLTEPILGFLFGYINRGSM
EYCFNCAQSSLGET

CPn_0554 637806 638141
CT440 hypothetical protein
VFSYLLLCIILVYVRFMYEGKSRMASPTPGQLHLQQKVESKAYDYSRSLAMIATALLFFI
VALILSGLSLLPQVFLPFSGAYFIIGSFLAFIALGILLINCVCDLKQYLTSS

638298 640241 CPn_0555 638298 640241

tsp-Tail-Specific Protease

MFVMKKLVRLCVVLLSLLPNVLFSSDLLREGIKKMMDKLIEYHVDAQEVSTDILSRSLS

SYIQSFDPHKSYLSNQEVAVFLQSPETKKRLLKNYKAGNFAIVRNINQLIHESILRARQW
RNEWYKNFKELVLEASSYQISKQPMQWSKSLDEVKQRQALLLSYLSLHLAGASSSRYEG

KEEQLAALCLRQIENHENYYLGINDHGVAMDRDEEAYQFHIRVVKALAHSLDAHTAYFSK

DEALAMRIQLEKGMGGIGVVLKEDIDGVVVREIIPGGPAAKSGDLQLGDIIYRVDGKDIE
HLSFRGVLDCLRGGHGSTVVLDIHRGESDHTIALRREKILLEDRRVDVSYEPYGDGVIGK HLSFROYLDCLRGGHSTYVUDTHRGESDHTTALKRERT LLEDRKVUSTEFTOSYTAK
YTUHSFYEGENQVSSEDOLRRATQGLKEKNLLGLYLDIRENTGGFLSQATKVSGLFMTNG
VVVVSRYADGTMKCYRTVSPKKFYDGPLATLVSKSSASAAETVAQTLQDYGVALVVGDEQ
TYGKGTIQHQTTTGDASQDDEFKYTVGKYYSPSGKSTOLQVKSDILIPSLYAEDRLGER
FLEHPLPADCCDNVLHDPLTDLDTGTRFMFGKYYLPNLQKQETLWREMLPQLTKNSEQRL
SENSNFQAFLSQIKSSEKTDLSYGSNDLQLEESINILKDMILLQQCRK

640921 CPA_055KD: Cysteine-Rich Protein
ENGMSSNLHPVGGTGTGAAAPESVLNIVEETAASGSVTAGLQAITSGFGMVNLLIGWAKT
KFIQPIPESKLFQSRACQITLLVVGILLVVAGLACMFIFHSQLGANAFWLIIPAAIGLIK
LLVTSLCFDEA-TSEKLMYFQKWAGVLEDQLDDGILNNSNKIFGHVKTEGNTSRATTPVL
NDGRGTPVLGFLVGKIARV

CPI_0557 642870 441144

omcB=66kDi_dysteide=frich_6MP
EIFMEKLIRRVTVLAUTSMACGFACWITEAAVAEGLITKIVAGAETKPAPVEMTAKKVR
EIFMEKLIRRVTVLAUTSMACGFACWITEAAVAEGLITKIVAGAETKPAPVEMTAKKVR
LVRRINKGPVECKSRESAFCDKEFVI-CHEGYRCQUPEAQQEGCYGRLYSVKVMDDX:NVETCQS
VPEYATVGSPYT_LEILA-GKKDC/DW/TTQQLPCEAEFVGODFTTTTTIDGKLWK/LDRL
GAGDKK/LTWAKTELKEAKSCFTAATVCAGPELRCYTK/CQPALCIRQEGPORTEXIEDMKPWATTTYVADIKGGACTYCHIGAGPQULSFNLGDMRFJOKKVETVEFCTPGRRG
QFTMVATTTYVAGBRGGANVTPV/HER/SQVMICHADMSYVCKPVEYSTC/VGNYDDLYLII
DVVTQFTPS-NTTVLEATSGETC/GRV/WRITKEMGFGETLQFKLWVRAQVI-GRFTNQVI-GESNGC/TCTSCAEFTPTIMK-GLAATHMC/YLDTNIDD-CVCENTVYRICVTNR-GBAEDTIMG
LILKPS/KELQPTAGGFTKSTCORTTY/PPDALPRUOKEGVEFOVTLKG-CATSDARGEAT

LEGOTETHEVEDTENTHY

543031 64 1 106 CPR_USSA - OVSTOINE-RICH LIPOPTOTEIN
KLMKKAVLTAMPCGVVSLSSCRIVDCCFEDPCAPSSCNPCEVIRKKERSCGGNACGSY
VPSCSNPCGSTECNSOSPOVKGCTSPDGRCKQ

EPRESENT BACKER 543927
TERRELE BYSIGHT STATE COLORS OF THE PROPERTY LEACHER LANGUE FLANCIST SELECTION OF THE PROPERTY SELE

645666 644098

CPn_0560 645666 644098

qltx-Clutamyl-trna Synthetase
RNSRFQQMKSLWSKDKRIMWENVRVRVAPSPTGDPHVGTAYMALFNEIFAKRFKGKMIL
RIEDTDRTRSRQDYEBIFSALRHCGIOMDEDPDVCGPYGPYRGSERTKIYGGYVETLLK
TDCAYKCFATPQELAEMRAVASTLGYRGGYDRRYRYLSPEEVASREAAGQPYTIRLKVPL
SGECVFEDYSKGRVVFPWADVDDQVLVKSDGFPTYHFANVIDDHLMGITHVLRGEEMLSS
TPKHLLLYEAFGWEPPYELHMPLLLMPDCTKLSKRKNFTSIFYRDSGYVKAEFVNFFIL
MGYSMEGDEEVYSLERIIETFNPRRIGKSGAVFDIOKLDMNKHYLNHEGSPECLLKELQ
GWLLNDEFFLKILPLCGSRITTLAEFINLTSFFFSGLLEYRVEELLPQALSPEKAAILLY
SYVKYLEKTDQMTKETCYLGSKWLAQAFMYHKKKAIIPLLVAITGKKQGLPLFDSIELL
GKPRARARLVYAEKLLGGVPKKLAATVDKFMQREDFEEATFDL

646407 645871

CPn_0561 646407 645871
euo-CHLPS Euo Protein
LMACEOHEGCYELEEREEIEDIKOSDTKWVSITQAAKLHNVTRQAIYVAIKQKKLKASKE
LMACEOHEGCYELEEREEIEDIKOSDTKWVSITQAAKLHNVTRQAIYVAIKQKKLKASKE
TRWEIDIKDLEEYKRNRYSRKKSLYQGELVFDNGKGCYSINQVAQILGIFVQKVYYATRT
GTIRGERKGAAWVIHVSEIERYKNEYLSKQAAKKLKGAEPKEHQAPNFEPPTEIFPESN

CPn_0562 648051 646918

*CHLPS 43 kDa protein homolog_1
NYKVIMSIAIAREQYAAILDMHEKPSIAMFSSEQARTSWEKRQAHPYLYRLLEIIWGVVK
FLIGLIFFFIPLGLFWVLQKICQNFILLGAGGWIFRPICRDSNLLRQAYAARLFSASFOUN
VSSVRRVCLQYDEVFIDGLELRLPNAKPDRWMLISNGNSDCLEYRTVLQGEEVDHIFRIAE
ESQSNLLIFNYPGVMKSQGNITRNNVKSYQACVRYLRDEPAGPQARQIVAYGYSLGASV
QAEALSKEIADGSDSVRWFVVKDRGARSTGAVAKGFIGSLGVWLANLTHWNINSEKRSKD
LHCPELFIYGKDSQGNLIGDGLFKKETCFAAPFLDPKNLEECSGKKIPVAQTGLRHDHIL
SDDVXKEVAGHIQRHFDN

CPH_0563 650113 648293

rec_i_ssdna exonuclease
OYMMLLMDFSPKGPCGIKFMINSDNASAAGLLWAHPKEDPAFLGMIIKEFHLPPTVAQIF
ISRGFQTIQEIHKFLYSHLSSLYDPGLFLDMSKAVERLLLARDRKEHVMIYGDSDVDGMT
ISRGFQTIQEIHKFLYSHLSSLYDPGLFLDMSKAVERLLLARDRKEHVMIYGDSDVDGMT
GVÄLLVEFIRDIDVMVSYFFLGAILKQHGETSTLIAKLKEEGITLLITVDGJIRGKEVS
DIFRGGIDVIITDHHMPTGKIPHCVATLNPKLDMTYSPRELIGVVAPKLARGVINALI
SRNIVPKSQSLKKLLDLVTLGTITDVGVLLGENRVMVRYGIKELARGARFGLNKLCALC
GVEKSEVTSTDIVLKIAPKINSLGRLDDPAKGVELLLTQDDERVDALIMELDNINRERQR
IEÄSVFQDVQEILNSNPEILKQAAIVLSSTAMHARVIPIISARAKTYNKFVVIIAIQRG
IGKSSARTIGSFPLLGVLKKCSSLLLSYGGHDFAAGVINKEDKVEDFKKKFVHLNNSSLK
KGPTLPHLEIDAYADFDAIDVDLLASMELFEPFGKGNLMPIFYSKVRQVRYPKVLPGMLL
KLYLSQKERNLEGVAFGLGRHADALKASWHYPLEIAYTPRLSOTSGSGVIHLLVRDFRIS
SEPPRISD

CPB_0564

CPB_0564

CPB_0564

GSAKKOKVKRNFAIIICVFALALYYVLPTCLYYAKPLDKKIDONEAHHIKSFTKOAOOV

SGAKKOKVKRNFAIIICVFALALYYVLPTCLYYAKPLDKKIDONEAHHIKSFTKOAOOV

RKÖVIPRVSAILSSLHLRGHIQOHPAIPDIVSVFRKGEDAEDFIONLYHGEPNYPIKSA

KRÖVIPRVSAILSSLHLRGHIQOHPAIPDIVSVFRKGEDAEDFIONLYHGEPNYPIKSA

KRÖVIPRVSAILSSLHLRGHIQOHPAIPDIVSVFRKGEDAEDFIONLYHGEPNYPIKSA

KRÉGLTLHSDLLAQRTSLSKEQRLDFDSFVSYSSENEQEMASSILORVYSACFFPK

LLHMAGYSREHDDHVIQVASSINTSLVESDFSFVSYSSENEQEMASSILORVYSACFFPK

SLÄMDAIDVEDQKLLKSVYETLSOTAGIRSLORPYTEGKRLDCSESSLFFSSIFYCPKE

SLÄMDAIDVEDQKLLKSVYETLSOTAGIRSLORPYTEGKRLDCSESSLFFSSIFYCPKE

SKÄELTLHSDLLAQRTSLSKEQRLDFDSRLAVEKOKLSKKLTVQVEDYNNGFSFOMMOKD

KRÍELTLHSDLLAQRTSLSKEQRLDFDSRLAVEKOKLSKKLTVQVEDYNNGFSFOMMOKD

KRÍELTLHSDLLAQRTSLSKEQRLDFDSRLAVEKOKLSKKLTVQVEDYNNGFSFOMMOKD

TNDŽÄHFSKGSVYILLKGLRSIVAKYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NDÖŽÄHFSKGSVYILKGLRSIVAKYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NDÖŽÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NTDČÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NTDČÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NTDČÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NTDČÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNCFSHTEAISWTLGE

NTDČÄHFSKGSVYILKGLRSIVATYQOGGKELOSFFKDLQNLYNGFROFQVILSFKL

MSFHVVNERFSSYSASRYEVORFLDYLMFTSOAOCKTOSPESITYTEASALFNEEVDVPSV

MSFHVVNERFSSYSASRYEVORFLDYLFTSOAOCKTSPESITYTEASALFNEEVDVPSV

MSFHVVNERFSSYSASRYEVORFLDYTFSTSOAOCKTSPESITYTFASALFNEEVDVPSVAN

MSFHVVNERFSSYSASRYEVORFLDYTFSTSHTWTSAVCOPGISGTANOQYSAN

KOIRBEFAAGEGYVLNSVKUTSKVTSVKSKKKKKSVEKSVEKKSVEKGYTKAFGAIFDSNLTT

VLASALLFFLDTGFIKGFALTLLIGIFSSMFTALFMTKFFFMLMNTQHTQLHANNKTV

VLASALLFFLDTGFIKGFALTLLIGIFSSMFTALFMTKFFFMLMNTQHTQLHANNKTV

VLASALLFFLDTGFIKGFALTLLIGIFSSKTTALFMTKFFFMLMNTQHTQLHANNKT

VLASALLFFLDTGFIKGFALTLLIGIFSSKTYMTTATTLSVLLMLLYFSDKALSYTRADTSLSPKINDHELL

AVGLLSETGLDFSTETLNETONFÖKKVSKLKKKMPYOATLGLLGALAIILLVSLFFEW

OYAFSAVCALIHDLLATAVLFIAHFFLKKIQIDLQAIGAMTVLGYSLNNTLIIFDRIR

EDROANLFTPMHVLNNDALOKTERRKE LGTLSSLYIAPPLLLFMVRKENRSK

CPn_0565 655741 654533
CT449 hypothetical protein
NKLFCFLIFCFVNISAILFDSSFLLKIKRNSKRMLRSMKFPRISISDLIPTOMVIWMRGG
RNWHYVPNAQNLPKRILGGVLACFGLALLGCAAFAGGVCGTIFPCIGLMILGLVLLGFAY
LQVSKGWSRFERPLFRETKVFEKPIMMLGCLSLIGSWKKIRPGCYYHPGCPQVEICEGSG
EIVTKIFQKKDDRNTSIFLIQEMDQIALRQGIFKSSLSRKFFAIDFSVVSSLLSSIGNEE
GQYLDFKVISMSEGQASDBTHPKSAITYNISGAAGEPGGRCYIDAYTKAFFTVLDQIGD
INIVKHITIYVLTPILGVPDALPKEEQENLK/LLSQAAFLYSAEQVAKRMREEKQDSIRIK
FIFTDPTSFTCLYFSPHHSSTPHSVTPISLSGFVGEQESYTFA

650099 696890 Cfm_0566

CHALO908
WHICH THAN THE PROPERTY OF THE PROPER

657817 cdmA Phospharidate Cyfidylyrransterase VLNSNKFKSKTSAYODLFG

LVLTELVLLLYSSLFFLTSFALGFLTATCGAVGTY
EYSSMAKAKMYPLSTFSALG

FLALSFLJTRWGHSLOGFFDALDATLLLVWVWSIF
RVRKSTIGALOLGSVTLFSILVYGIFIRLFLAVLYSFHTDGEPYLJTWAGFLIATTKGA
DIFGYFFGKAPGNKKIAPQISPNKTVVGFVACCLGATLISFIFFLQIPTRFASYFPMPAI
LIPLGLALGITGFFGDIIESIFKRDAHLKMSNKLKAVOGMLDTLDSLLLSTPIAYLFLLI
TOSKFFIG TOSKEFIG

CHA PROGRAM FROM YEAR OF THE THE TOTAL OF TH 657805 658464

658,898 659099 CPn_0569 658,98 659099
plsc-glycerol-3-P Acyltransferase
Legfdnktssgenfsftiskramifrickfftwvafslfyklkvygvkknfikgpaiiav
Legfdnktssgenfsftiskramifrickfftwvafslfyklkvygvkknfikgpaiiav
NHNSFLDPIALHMCVHEGYHLARASLFNIPWLMKQMGCFPVRQDEGNSAAFKIASRLFN
NHNSFLDPIALHMCVHEGYHLARASLFNIPWLMKQMGCFPVRQDEGNSAAFKIASRLFN
KRKKLVIYPEGARSPDGGLOPGKVGIGMMAAKSRVPIIPVYIRGTFEAFNRHQKIPHVWK
TITCVFGTPMYFDDIIGNPEIKNKETYQIITNQTMNKIAELKAWYESGCKGDVP

CPn_0570 659044 660789

args-arginyl tina transferase

TKLPSSKHGNNGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

TKLPSSKHGNNGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

TKLPSSKHGNNGGKETSFKLMSTLLSILSVICSQAIAKAFPNLEDMAPEITPSTKEHFG

TKVCHDAMKLARVLKKAPRAIAEAIVAELPOEPFSLIEIAGAGFINFTFSPVEHDVLRINH

KVDALKIGFOVSQPKKIIIDFSSPNIAKDMYVGHLRSTIIODSLARIFSVVGHDVLRINH

IGDMATAFGHLTYLQENPCDYSDLEDLTSLYKKAVVCFTNDEFFKKRSQONVALQAKD

ROMATAFGHLTYLQENPCDYSDLEDLTSLYKKAVVCFTNDEFFKKRSQONVALQAKD

KOVFHEAFS/PFNVQKSDGSYNYATTDLAAMFYRIEDHADKIIIVYTDLGQSLHFQLLED

KCVFHEAFS/PFNVQKSDGSYNYATTDLAAMFYRIEDHADKIIIVYTDLGQSLHFQLLED

TAIAAGYLGFGIFSHVGFGLVLDPGGKKLKTRSGNVKLRELLDTAIEKAEEALREHPPE

LTDEAIOGRAPVIGINAIKYSDLSSHRTSDYVFSFEMLRFEGNTAMFLLYAYVRIGGIK

RRLGISQLSLEGPPEIQEPAEELLALTLLRFPEALESTIKELCPHFLTDYLYNLTHKFNG

FFRDSHJODSPYAKSRLFLCALAEQVLATGMHLLGLKTLERL 659044 660789

CPn_6571 662179 660749

murk-UDP-N-Acetylglucosamine Transferase

TF#XVNVSSSDFDAKGERRNJIAVYECGRINGEVKVSGAKNAATKLLVASLLSDOKCTL

RNVPDIGDVSLTVELCKSLGAHVSMOKETEVLETYTEEIQCTRVPPTFSNVNRIPILLLG

RNVPDIGDVSLTVELCKSLGAHVSMOKETEVLETYTEEIQCTRVPPTFSNVNRIPILLLG

MLIGRCPESVVYPTVGGDAIGERTLNFHFEGLKOLGVOISSDSSGYYAKAPRGLKGNYIH

ALIGRCPESVVYPTVGGDAIGERTLNFHFEGLKOLGVOISSDSGYYAKAPRGLKGNYIH

ALIGRCPESVVAFENLIAAHAAKGRTVIKNVALEALELIDLVLFLQKAGADITTDNDRTIDIFG

TGGLGSVDHTILPDKIEAASFGMAAVVSGGRYFVRNAKOELLIPFLKMLRSICGGFLVSE

SGIEFFGERPLVGGVVLETDVHPGFLTDMQQFFAVLLSQAGGSSVIHETVHENRLGYLMG

LQHMGAECOLFHQCLSTKACRYAIGNFPHSAVIHGATPLWASHLVIPDLRAGFRYVMAAL

LABCGGSIIENTHLLDRGYTNWVGKLRSLGAKIQIFDMEQEELTTSPKSLALRDASL

IABCGGSIIENTHLLDRGYTNWVGKLRSLGAKIQIFDMEQEELTTSPKSLALRDASL 660749

CPD_0572 662349 664616

CT456 hypothetical protein
IMAAPINOPSTTTOITOTOOTTTTTTVGSLGEHSVTTTGSGAAAOTSOTVTLIADHEMQE
IASQUGSAVSTSAEHSSTLPPETGSVGATAQSAQSAGLFSLSGRTORRDSEISSSSDGS
ISRTSSNASSGETSRAESSPDLGDLDSLSGERAEGAEGEGPGDPGGLPEST IPHYDPTDK
ASILNFLKNPAVQOKMOTKGGHFVYVDEARSSF IPVRNGDMSTAES KVSNAKTKENITK
ASILNFLKNPAVQOKMOTKGGHFVYVDEARSSF IPVRNGDMSTAES KVSNAKTKENITK
RADLEMCIAKRCVGYETIHSDWTGRVKPTMEERSGATONYHLMLSKKFKTAVVYGPMNA
KESSGYTPSAAMRGAKVETGP IWDDVGGLKGIMKKTTPAPDFSF INETFOGGAHSTSHT
KESSGYTPSAAMRGAKVETGP IWDDVGGLKGIMKTTPAPDFSF INETFOGGAHSTSHT
ISTSGOSTIEEDTIQFDDPGQGEDDNAIPGINLGGITTNVTTEGGGTNITSTKSTSTDKVS
GPGTPVGATVVPNVNNHLGGIKVDLGGIKNGUSVDFPTVILDFKTGDTDPSGQATGG
LQNVRQHLNTAYDSMGNSVSDLNQDLGQVVKNSEDAVMFPTVILDFKTGDTDPSGQATGG
VTDGGGHIRNIORNTGSTGQSEGATPTPOPTIAKIVTSLRKANVSSSSVLPOPOVATTI
TPOARTASTSTTSIGTGTESTSTTSTGTGTGSVSTQSTGVTPTTTTRSTGTSATTTTSS
ASTOTPQARLPSGTRHVATISLVRNAAGRSIVLQOGGRSQSFPIPPSGTGTQNMGAQLWA
AASOVASTLGOVVNQAATAGSQPSSRRSSPTSPRRK

664691 665413 CYPL_D5/3
yebC family
yebC family
VEMAGHSKWANTKHRKERADHKKGKIFSRIIKELISAVKLGGADPKSNARLRMVIQKAK
VEMAGHSKWANTKHRKERADHKKGKIFSRIIKELISAVKLGGADPKSNARLRMVIQKAK
ENNIPNENIERNLKKATSAEQKNJEEVTYELYGHGGVGIIVEANTDNKNRTASDHRIAIN
KRGGSLVERGSVIVNFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVERGSVIVNFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVERGSVIVNFARKGACTVAKSSIDEEVIFSYAIEAGAEDLDTEDEENFLVICAP
KRGGSLVERKLISQGATCSEDRLIYLPLRLVDCDEKDGEANLALIDWLEQIEDVDDVYHNM CPn_0573

CPn_0574 665978 665394
No robust homolog present in Genebank/EMBL as of 11/7/98
SAERGFRHBIVMVETVLHNFORYLSKYLYRVFRFPCRKKTFLSSHRVLARPSFPVDYCPG
KIYDLQEIYEELNAQLFQGALRLQIGWFGRKATRKGKSVVLGLFHENEQLIRIHRSLDRQ
EIPRFFMEYLVYHEMVHSVVPREYSLSGRSIFHGKKFKEYEQRFPLYDRAVAWEKANAYL LRGYKKRVGGGYGRA

CPn_0575 666524 665982
Yhhy-Amino Group Acetyl Transferase
SIFORWARSFMTAEKONTGILGLEIRYTLPSDATYMLKWLNDPKILRGFPIOTEAEIRET
VNFWVGFYRYHSSLTAVYNSNVAGVATLVLNFYVKVSHHALISIIVGEEFRNKGIGTALL
NNLIHLAKTRFKLEVLYLEVYEGNPALHLYQRFGFVEVGRQNRFYKDEIGYLAKTTMEKD

CPn_0576 667543 666494 prEB-Peptide Chain Release Factor 2 (natural UGA frame-shift) MQENLDKRLEALRTEISLAARSL

CPn_0576.1 p67598 prfB-(natural UGA trame-shift) MQENLDKRLEALRTEISLAAROL

CPN_0577 64786 207986 668155

POWITY AND COMPLEX BLOCKETT CONTRACTOR CONTR

CP0_0578 668388 mile 188 60.0365 CPG_0574

you'l-phosphohydroloss
TTHERETVLES SLATLE LLAPEWAGE FEIRMLEPTATEWRLPKKHARLISCER LAQUEDLAF
TTHERETVLES SLATLE LLAPEWAGE FEIRMLEPTATEWRLPKKHARLISCER LAQUEDLAF
THREVEEKER LIKEVISK SKREGEDLIVFFYSDLLGFARLEDKERLETFLATTLEADTAVFALL
CHILDYSOY I SRNTKGE LTC LPEEKISK! LOKAT LAVMOTLESS DYTRY DENUT OPPHEDL

99

EKLLKNTPUTLLHNTTHVIPNTLNIVGLGDLFARGFH LTRLQQYPDFVLSGHSHGPOVTLSWPKFARKFFERL NRGLGULKRIRFCSPPEICYITCSYD MYDPSLPGLLLSHNPDG PYLARGYFYTKEGKQLYV

CPn_0579 669310 669993
ygbP/yacM-Sugar Nucleotide Phosphorylase
KEFAGAPLLKGATSHVPHIKSSLILLSGGGCTRFGSKIPKQYLPLNGTPLVLHSLKILSS
KEFAGAPLLKGATSHVPHIKSSLILLSGGGCTRFGSKIPKQYLPLNGTPLVLHSLKILSS
KEPAGAPLUKGATSHVPHIKSSLILLSGGGCTRFGSKIPKQVLPLNGTPLHTLHDGARPFIY
LPGHIH LAISTARFI GATALAJETIYTIKORIPVPHIGPDNICATHTETJATTETLRFGLA
LAFRKGATTVDDTEAMELFGETQLVFRKHDGIKLTYPEDLFTAGALL

CPn_0580 669936 670793
truA-Pseudouridylate Synthase I
ASSNOMFLPRRSNDCPSPMTKVALLIAYQGTAYSGWQQPNDLSIQEVIESSLKKITKT
RTPLIASGRTDAGVHAYGQVAHFRAPDHPLFANANLTKKALNAILPKDIVIRDVALFDDN
FHARYLAIAKEYRYSLSRLAKPLPWQRHFCYTPRHPFSTELMQEGANLLIGTHDFASFAN
HGRDYNSTVRTIYTLDIVDWGDSLSIICRGNGFLYKMYRLVGALLDVGKGAYPPEHLLD
ILEQKNRREGPSAAPAYGLSLHHVCYSSPYNNFCCEQCSVSTSNEG

CPn_0581 671533 670745
Phosphoglycolate Phosphatase
EGLRWRSYKSTLRQCWIYSMLVSDEFQLCLRSGMYLEDYDVFFFDLDGLLVDTEPCFYRA
FLQACAEFSLEVHMDFSTYYSHTTLGTEIFSKKFIEQYPQAQEYMAEIFAKRLQIYYKSL
EHAGPALMEGVEAF IELVLSLNKTFGVVTNSPRDATHTLRTMYPILNKFLFWTREBYAR
PKPYGDSYDYAYRTFAREGMKVIGFEDSVKGLRALSKIPATLVCINSMAEITPEDYPELK
GKEFFSYPSFDVLTEHESQQKLL

CPn_0582 671305 672177
CT465 hypothetical protoky construction of the knewallkkiohrlukhdknkvlyloanhlmokrkrhnplntyhssmitetrrlptyk snivlkmilristvslltscsfsknsrtfvfyberitsokocpulhekstrispelydd ispnrevitaysfycrogonsiitpegvlydcdlhhsitreefryiherletwyrllog dhensitreefryiherletwyrllog dhensitreefryiherletwyrllog dhensitreefreiherscislsolhlogtaatfaldpelpmekllatikklykk nsdpslsnfivteatltnpelrltogoloshteitveildnlonkealssa

CPn_0583 672349 672717
CT466 hypothetical protein
ULSFFLCKTKVTPRFLMMERTLLLLLKKKKGLFLAILDLTQTESSLTTPELEKVLKQKK
IFLSCIDRVDLQIKEFRHAFSSELPQDIQEELEEIRDVIIRILDTDKRNYAQKKKEFGIY
ERP
####

CPH_0584 672659 673798

at65/ntr8-2-Component Sensor
IRTMATHHRKKRNLVFMWYDSNNLHPPAYELLEIKARITOSYKEASAILTAIPDGILLL
SETCHFLICKSOAREILGIDENLEILNRSFTDVLPDTCLGFSIQEALESLKVPKTLRLSL
CKÉSKEKEVELFIRNNEISGYLFIQIRDRSDYKQLENAIERYNNIAELGKMTATLAHEIR
NÉSGIVGFASILKKEISSPHORMLSSIISGTRSLANNLVSSMLEYTKSQPLNLKIINLQ
DEESSLIPLLSVSFFNCKFVREGAQPLFRSIDPDRMSVVMNLVKNAVETGNSPITLTLH
TSGBISVTNBGTIPSEIMDKLFTFFTTKREGNGLGLAEAQKIIRLHGGDIQLKTSDSAV
SEFIIIPELLAALPKERAAS

CHICOSS 675880 673865

STÄLIARITY TO CPS INCA,2

STÄLIARITY TO CPS INCA,2

STÄLIARINFSIGDCSSNMATPAQKSPTFQDPSFVRELGSNHPVFSPLTLEERGEMA IÄRVQCGMNHTIVKVSLITLALLTILGGGLLVGLLPAVPHFIGTGLIALGAVIFALALI LGLYDSOGLPELPPVPPPQOIOIEDLRHETEVLEGTLLEVLLKORDAKDPAVPQVVVD CEKRLGMLDRKLRREEEILYRSTAHLKDEERVEFILLELERSLVADRLEFNRRSYERFV GÄMTVRSEGEKEISKLQDLISLQQTVQDLRSRIDDEQKRCMTALQRINGSOKDIGNA HDREASGRACEGTEMDCAERQOLEKDLARQLKSMGEVIEHRGTIHQQEKAMRKONAKLER LØFFIRLTGIAPDEGSLETYREYVEKVLSOKLDMOKILQEVNAEKSEKACLESLVHDYEKQ LEDKDANLKKAAAVWEEELGKQOQEDYEGTOPIEIRLISTFILEVQDSLREAEKVEKDFQL LEDKDANLKKAAAVWEEELGKQOQEDYEGTOPIEIRLISTFILEVQDSLREAEKVEKDFQL OKNYSKLQEKOVKKK ILEESMNHFADLFEKAQKENMAYKKKLADLEGAAPTEIGEDDD WALTDSASLSOKKIRELVEENQELLKALAFKSNELTQLVADAVPAEKEISKLREHIEDOK GESTÄALDKHAQAIKOCEAAQRKCCDLESLLSPVREDAGMRFELEVELQRLQEENAQLRA EVERLEQEGFQG

CPTED586 675993 677183
acC/ntrc-2-Component Regulator
KEKMNPSRCENMAIKNILVVDDEPLLRDFLSELLTSQGFIPDTAENLRNALQMIRSRDYD
KEKMNPSRCENMAIKNILVVDDEPLLRDFLSELLTSQGFIPDTAENLRNALQMIRSRDYD
KUYISDMSMPDGSGLDLIKI IKQSSPHTPVLVVTAYGSIENAVEAMHQGAFNYLTKPFSSF
ALFAFISKAEELKNLVHENLFLHSQTTPDSHPLIAESKAMKDLLAIAKKAASSSANIFIK
GESGCCKEVLSFFIHHNSPRANHPYIKVNCAAIPETLLESELFGHEKGAFTGATTKKAQR
FELAHKGTLLDEITEVPVNLQAKLLRAIQEKEIEHLGGTKTLSVDVRILATSNRKLKA
IDDKSFRQDLYYRLNVIPLHLPPLRDRQDDILPLANYFLNKFCRMNNTPLKTLSPKAGEL
LLNYPWPGNIRELSNVLERVVILENTSLLTEDMLALA

CPn_0587 677378 678124

*YVYD_Bs conserved hypothetical protein
SYGELFILSTLIKHHVTLGDKMRPHRKHVSSKSLALKQSASTHVEITTKAFRLSYPLKQL
ILEKSDH_DPMETIRVVLTSHKDKLGTEVHVVASRGKEILQTKVHNANPYTAVIMAFKKI
RTHANKHSNKRKDRTKHDLGLAAKEERIAIQEEQEDRLSNEWLPVEGLDAWDSIATTLGYV
PASAKKKISKKMSIRMLSQDEAIRQLESAAENFLIFLNEQEHKIQCIYKKHDSNYVLIE
PSILKPGFGT

CPn_0588 678033 678626
CT469 hypothetical protein
TSKSIKSNAFIKNMTATMSLLNLPSSQDSASEDSTSQSQIFDPIRNRELVSTPEEKVRQR
LL:SFLMHKLNYPKKLI IIEKELKTLFPLLMPKGTLIPKRRPDILI ITPFTTTDAQCNTHN
LGDPKPLLLIECKALAVNONALKQLLSYNYSIGATCIAMAGKHSQVSALFNPKTQTLDFY
IMJEEYSQLLNYFISLNL

CPG_0589 678634 679395
CPG70 hypothetical protein
CBG0 (CYTYO hypothetical protein
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPEGLETFFAKGGGTLQCDYRETLVP (GLGKYT
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPEGLETPT
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPEGLETPT
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPT
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPT
CBG0 (CYTYOVLRGRPLIKNIFFLTPLETPT)
CBG0 (CYTYOVLRGRPLIKNIFFLTPT)
CBG0 (C

Cfm_0590 680136 679516 CT471 hypothetical protein LFLYGDHNLGFACRYLFFF GASFUNGLLGVPTWLHEBERFYTHREESKSYPDME NMEIQAQRKKRVEFILTGEFALL TLMYQSGFGHLRAKCRUVYPVLYALNFSCSSCKMDM DFRGKWNRSSTITISNOKESINÚKLEKDVGVIVNTKTSLKGNVCPGSTFIKGSWGVWNKI YHNDLVGFSEVTLIFNVSSECGTITFS

CPn_0591 680 64 681020

CPH_D591
yage family
sLimrCTAYCTASAYNLHVU_FHLLKPRYPTILISREYVLANLDSTQASNQLAIFFPFGVAV
SLIMRCTAYCTASAYNLHVU_FHLLKPRYPTILISREYVLANLDSTQASNQLAIFFPFGVAV
LORINGAYNLHUTHTTITTITY
LORINGAYNLHUTHTITTITY
LORINGAYNLHUTHTITTITY
LORINGAYNLHUTHTITTITY
LORINGAYNLHUTHTITY
LORINGAYNLHUTHTY
LORINGAYNLHUTH

CPn_0592 /681132 681461 yidd family LYSKMFSMSFKRFLOO PVRICLLIIYLYOWLISPLLGSCCRFFPSCSHYAEQALKSHGF LMCCWLSIKRIGKCGFMHPGGIDMVPKTALQEVLEPYQEIDGDSSHFSE

CPn_0593 682494 681391
CT474 hypothet cal protein
VLGAKCMAFKRITISMLMOVILISVGLINHLFLLLFYSAIFRKDIYKLHLFSGFLIAKSSRK
VYLSEDFLUBISOÁSLDDLISLFKDERYMYGRPIKLMALSVAIASHHIDITPVLSKPLTY
TELKGSSVRWLLBNIDLKDFPVILDVLRCHKYPYTSKGLFLLIEKMYDEGWYDEDCLYHF
CSTBEFLYLRTLLVGADVQASSVASLARMVIRCGSERFFHECHEESRTSMISATOROKVL
KSYLDCESLAA/LLLUVHDSDVVLHFFCDEDLEKVIRLMPQESPYSONFFSRLQHSPRRE
LACHSTORVEAPRVQEDQDEEYWQDGDSLWLIAKRFGIPMDKIIQKNGLNHHRLFFGKV
LKLPAKQS

CPn_0594 682517 684958

phet-phenylalanyl trna Synthetase Beta
ntchytovivkslvktslrlssmripitlldtyfseplstkeileacdhigieaeientt
lysfasyltakilhtiphenanklrvatltdgekehowccapnceaglivalalegakl
ptsecoaytikkslkdsvesogmccgadelgldelqigerallelpeatplgedlatvlg
ntsleisltphlghcasfiglareichvtoanlvipkefsfenlpttaldmondpdicpt
fsyvvifgisaoppiklogsloalkokpinaivbitnyihkisgoplhaydashvalds
fsyvvifgisaoppiklogsloalkokpinaivbitnyihkisgoplhaydashvalds
fsyvvifgisaoppiklogsloalkokpinaivbitnyihkisgoplhaydashvalds
fsyvvifgisaoppiklogstoplkokpinaivbitnyihkisgoplhaydashvalds
geicrelkevalrpktlogispisipitnonvvaloaaihyileifpeatispiyss
geicrelkevalrpktlorilcksfsiellsokolgfstropetsllvkybsyrhdin
eldveeicrtesmaietonpvscytpiyklkretagflanagloeffpolldpetva
ltrwkkeeislogskhttvlrssllpglkksaatnlnroapsvoaffigtvyahdgopy
ocytlailltddessawlpkpslsfyslkowerellyhhhlsidaltlessalcefhpy
ocytlailltddessawlpkpslsfyslkowerellyhhhlsidaltlessalcefhpy
ocytlikkosfatlgovhpelakkaqikhpyffeelnldllcomlkkttklyryaiyp
ssfroltitveedipanllroklhegskwlesvtiisiyodksletrnknvslrlvfod
ybrilsnodieeeycrlvallnelltdtrgtins

DEN 0595 684943 685926
TT4TK hypothetical protein
RDYOFKYOLLFCVCYBANGCSAYASPRRODPSVMKETFRNNYGIIVSCOEWVKRGSDGTI
RDYOFKYOLLFCVCYBANGCSAYASPRRODPSVMKETFRNNYGIIVSCOEWVKRGSDGTI
RVVLKNGARTHEVYSCGLLHGEITLTFPHTTALDVVQIYDOGRLVSRKTFFVNGLPSOEB
LFNEDGFFVLTRWPDNNDSDTITKPYFIETTYCGHVIEGSYTSFNGKYSSSIHNGEGYRS
VFSSNJILLSEETFNEGWAVYTTFFYPNDPESITHYCNGOPHGLRLTYLQGGIPNTIEE
WRYOFOGTTIVFKNGCKTSEIAYVKGVKEGLELRYNEQEIVAEEVSWRNDFLHGERKIY
AGGYOKHEWYYRGRSVSKAKFERLNAAG

CPn_0596 685930 686457
ada-methyltransferase
FAVMADDTLIPKLMKNSLSQACSEGLLIAKYPPLQVIVHFDNNLVVKTHLSVAPVFSCLF
LGPAAKKAMGEIVLMCSRYANKEHPPFSSHFAKDLIPSQYLEILNCVAEIPFGEQQTYAE
IAKKTDTHPRTVGAACKQNPFLLFFPCHRVVGSHGERNYVLGPVIHEILLKFENSY

CPn_0597 688215 686479

oppc-01igopeptide Permease
MKHPSFYORFLSAYYKNLLASLSWKFFISVALIGIYAPLFASSKPLLVTWHGEIFFPLL
RYLFFPGYYTKPVDLFFDVLMVTFPFFILSFKLTRGWLRRWLLGLCIISOCMIFAWAYSG
KYOPALAENLKWRAEKVRENISKVNSEWWALLPKDTRTWEMERRYWSTYEQLGILIKA
KYRKKQEASVKKYQVAFEEKROSPMPTLRHLEMKNEGICLKRLQORVDKWGRPYEMAQOA
WNRATDNYRPFLMALTIEHELRLADYNNWGOPEDLCIAYANVEKRAEPYKKSLLEIRQV
LEDYAKLRSAISFIODKRLWIEKESEDLRILINPFFSSFHWEDDAGGSRENNKYVPWWQL
SRVTRKDLLAALVFGIRIALVVAGIGITIALAIGIMIGLVSGYFGGTYDMILSFFTEIWE
THPVLFILMUVISTTOQKSLLLNTVLLGGFSWTGFSRYVRIEVLKQRORGYVLAATNLGY
SHYYIMVHQILPNAIVPVISLVPFAMMAMISCEAGLTFLGLGEESSASWONLMREGVTGF
PAESAVLWPPAIILTMLLIAIALIGDGVRDALDPRLQDS

CPn_0598 689712 688219

ODDB-01igopeptide Permease
EEGGSVLKYILKRLVLIPLIPLITEAIVSINFVILNAAPGDVLEEKSRDALGEAGKSDKMRSY
EKGPDRYLGFREHYGLTLPIFFNTRPKITHKKIOTALGELANAINTTPSAKNAAKSLVYMG
DCAKFVMPALLFEADDASRDDKYRHITADLFIRGVLQGFVGPNLSPEGRAQNKEIAESN
AFLVRQLNEEDLDTNYEALKGWFOOMGGTEVFCVSSKOFWKTFFLETRFARYMSRVLRLD
FGTLRNDAHKTVISEVIKRLRCSLVLSILPHIVGFVLCQIFGMIMALKRNRWIDHSLNFI
FLILFSIPVFVAVFNILDNFVINKTIPFTTIPMPYSGLRSPPEVFNELSTLGRIFDLVSH
GFLPFCAVSYGALAAQSBLSRSIFLEEVLSQDFICAAKARGLRNFDILLYKHVGKNAAVSIV
TSLASSLGTLLGGALVVETLFNIDGFGNFFYQAILNRDHNVVLFSVLVGSALSLVGYLLG
DICYYLLDPRVQLEGRRI

CPn_0599 691823 589682
oppA-oligopeptide Binding Lipoprotein
kreschmykrovloktikkofvagslillyykssollerdiksikgnyrdigedireisry
vkovotsoaipaaniymlarklyrdeafallegobsyphilssldpykgottepellgtnyh
engliktaniygkpenlspengflyvgefydlipelalphydkgefesgoblaykieehly
engliktaniygkpenlspengflyvgefydlipelalphydkgefesgoblaykieehly
engloksperhiylrenyfwreidkrlywwkahpythedkkerrythidikffydaymmpyy
atmaavalricyeptvydysyrbiblikuvywkahpythedkkerrythyddikffydaymmpiy
atmaavalricyeptvydysyrbiblikuvywkahpythedkkerrythyddikffydayhphyddekivf
prevygyfangekileden idtyfphischanneitydogayyfamddekily
prevygyfangekileden idtyfphischen idty

CPn_9600 c02156 cd827 **L97827**No robust homolog present in Genelank/FMMEL as of 1177798

HGYMKIKKSFOYSLCOAKRFONMLPNHFDPCLOPVNU FSSLLKEETCSLNRAKOHLLYKILRDFNTMOHLRSLG

LAYGELIILLSKYQQKT

693092 592736

CT483 hypothetical protein
OFFRIHADDIINSMDEITPNYPLLRQDSLWNRVRVSWRADLSVSSRYEIASAIAILGLLV
AFCASAAVSIIFTANPLAQVFIDGCLALGLLPIPLVIGLLIIGIIVLLYGIYLFPQQRE

GITHT WHIM GAZIOY 1121_11401 602

TITIAL HYPOTHAL HALL PROTEIN CASTON (ASTON)
DSGFMKPLGFQENLEALCHKTSRQLLKYLIKQILFVCGASLLIALEFSFFLYFFLFSGKT
VIPAFCLACFFLTLFVCLVTRLYLLSGKCDFFEDLASEYLQGAVPPNKRSQNIVEEQSHL
AAAATKLSINLQNQEYSLLSEIFKFLPKHDLIRKFSCFCFWKDYFLFRECLLQKAIEAYI
KVVQAI FUDLSAHVSLADAYVALSGLYADPRKYPEFDANYWIPSGRYSAEIQEKFFATAR
RAIEEFQILNEYAPCNAWVHAQLAYSYHDLQMPMEEIQEYEIVLKLKPNDVETMSKLGIL
YFQQGMAKGLRIYEEIKKRDYKKSQKLIKFYGVEYKY

694136 695185 CPH_0603 694136 695185
hemZ-Ferrochetalase
wKimRilvLMQCLVSLFLAKKVTVTTPAYLLANFGGPRHAKDLQEFLISLLTDRDVTGTF
LPRVLHRHLFTFIAKKRVPKVLPOYQSLQNMSPIYFDTETLAKTLSEILRAPVIPFHRYL
PSTHEKTLLALRTLHTRHVIGIPLFPHFTYSVTGSIVRFFMKHVPEIPISWIPQFGSDSK
FVSLITCHIRDFLQKLGILEKECCFLFSVHGLPVRYISQGDPYSKQCYESFSAITTNFKQ
SENFLCFQSKFGPGKWLSPSTAQLCQNIDTDKPNVIVVPFGFISDHLETLYEIERDYLPL
LRSRGYRALRIPAIYSSPLWVSTLVDIVKENSTVVAEELIKSGKKHTGIR

CPn_0604 695981 695196
fliy-Glutamine Binding Protein
CKKRQNSEAQLAVKIKFSWKVNFLICILLAVGLIFFGCSRVKREVLVGRDATWFPKQFGIY
TSDTNAFLMDLVSEINYKENLNINIVNQDWVHLFENLDDKKTQGAFTSVLPTLEMLEHYQ
FSDPILLTGPVLVVAQDSPYQSIEDLKGRLIGVYKFDSSVLVAQNIPDAVISLYQHVPIA
LEALTSNCVPALLAPVIEVTALIETAYKGRLKIISKPLNADGLRLAILKGTNGDLLEGFN
AGLVKTRRSGKYDAIKQRYRLP

696737 696150

CPI_0003
yhbF-Methylase
LRKLCSSRGDVRILAGKYKGKSLKTFSNPHIRPTSGLVKEAFFSICREDIEGAAFLDLFA
GMGAIGFEALSRGASVVFVDISIKAIQLIHTNSALLGEQLPVVIFRQDAQSAIQRLIKQ
KRSFDLIYIDPPYELCNCYVETLLQKIVSGNILNPEGTLFLENASDEEIACEGLTLRRRR

c₽n<u></u>0606

CPH20606 697492 696707
CT#88 hypothetical protein
SSYSROLRFYTGSLOMHIYGLADLHLALGVPEKTMEVFGDFWIGYHOKICSEWQAVVHP
EDTVLLEGDISWAMMISEAHKOFAF IGOLPGTKYMIRGNHDYWSSASTSKILQALPFSLY
YLWGGFALLTPHLAVVGVRLWDSPTICVKKENFLTPSTQEQSYTEQDEKIFLRELGRLKR ARAALPKEVTEVIVMTHYPPISSDGTPGPISEFLEADGRVSLCLFGHIHKVQRPIDGFGN IRGIHYILVAADYVNFVPQEVM

CR_0607 698910 697573
gfgC-Glucose-1-P Adenyltransferase
NRBIOMIENDFPEASNFESSHFYRDKVGVIILCGEGKRLSPLTNCRCKPTVSFGGRYKL
IDIFISHAISAGFSKIFVIGOYLTYTLQQHLFKTYFYHGVLQDQIHLLAPEARGGDQIWY
OGFADAIRKNLLYFEDTEIEYPLILSGOQLYMMDFRSIVDTAIRTHVDMVLVAQPIPEKD
AYRKVVLDIDSEGKLIDFYEKPOEKEVLKRFQLSSEDRRIHKLTEDSGDFLGSMGIYLFR
RDSLFSLLREEGNDPSKHLIOAQKKRGQVOTLLYMGWADIGTIESYYEANIALTOKPH
AEKRGLNCYDDNGMIYSKNHHLPGAIITDSMISSSLLCEGCVINTSHVSRSVLGIRSKIG
ENSWYDQSIIMGNARYGSPSMPSLGIGKDCEIRKAIIDENCCIGMGVKLQNLKGYIRYDS
PÖKKLFYRDNIIIVPOGTHIPDNYIF PDKKLFVRDNIIIVPOGTHIPDNYIF

CED_0608 699690 699016

*UFFIGINE 5'-MONOPHOSPHATE SYNTHASE (UMD SYNTHASE)-Truncated?

VSFLYFYKNORRLWARMANYEDAKLROQAVAILYQIGAIKFGKHILASGEETPLYYDMRLV
ISSPEVLQTVATLIWRLRPSFNSSLLCGVPYTALTLATSISLKYNIPMVLRKELQNVDP
SDAKKVEGLFTPQOTCLVINDMYSSGKSIIETAVALEENGLVVREALVFLDRRKEACOPI
GRÖGIKVSSVFTVPTLIKALIAYGKLSSGDLTLANKISEILEIES

699986

CYPEUGUS 6396/2 639966 CT490 hypothetical protein QNTKNSLIRENMLIRLFLGISLPKGFPLYLEPPLVLATFQGTQFVGTYSEATNPLYJONL NLMYHYTQELLYKAVPCNYKSIYREIPLIIFPEVLIGSTPTQSTE

CPn_0610 701450 700029 rho-Transcription Termination Factor tho-Transcription Termination Factor
RIFILFROSIMKEERSSELLPRVKETKKHAYVSMOEKSCVGECAVVASESEEARSVTVTK
IAKLORMGIEELNILARQYGVKNIGSLTKSQVVFEIVKAKSERPDELLIGEGVLEVLPDG
FGFLRSFTYNYLPSAEDITVSPAQIRRFDLKKODTIIOTIRSPKEKEKYFALLKVDKING
STPDKAKERVLFENLTPLYPNORIVMEMGKDHLAERVLDLTAPIGKGORGL VAPPRSGK
TVILOSIAHAIAVNNPDIVLIVLLIDERPEEVTDMIRQVRGEVVASTTDEOPERHIQVAE
MVIEKARRLVEHGNDVVILLDSITRLARAYNTVQPHSGKILTGGVDASALHKPKRFFGAA
RNIEGGGSLTILATALIDTGSKMDEVIFFEFKGTONMELVLDRRLSDRRTYPAIDLIKSG
TRKEELLYHPSELERVYLFRQAIADLTTIDAMHLLLGRLKKTNSNAEFLLSLKE

CPn_0611 702133 701420
yacE-predicted phosphatase/kinase
RRNRRDAKTSEREDGISYDFIRSYSCEYLNWKLGRMLKLLKVSITGDLSSGKTEACOVF
OELGAYVVSADEISHSFLIPHTRIGRRVIDLLGSDVVVDGAFDAQA/IAAKVFYNSVLLQG
LEAILHPEVCRIIEECYHOSIODGNYPLFVAEVPLLYEIHYAKWFØSVILWMANEDIRRE
RFMKKTGRSSEDFDQRCSRFLNVEEKLAQADVVVENNGTKKELHGKIEEYFYALKGAL

704688 702022 polA-DNA Polymerase I pola-joha Polymetase t
RCHITCLLOTVERPRREYAMKKLEVLDASGETERAY FALPEMKNIQOATOAVEGETISL
NKLEKEESPEYM 13VEDGENNKOSRQA TYADYKSMEQKEEBTEPOTALVKEYYSELIGLA
YLERESVEADDV TASTAKKAREENYKVYVCTADKDLLOLVVDHVVANMPWADQGVVGTGE
VLERYGTERYM 14PCMLDAVGDSODN FRGLEVCQCRKKAAALLKOPGSVEGLLENLDAVKGL
SCHMLDEROETEKLISKILALLOSN FETEVETESLTP POHYVDEEKLHEYYCQGFRTLVP
SKOTEAATYDVO FEKDALELLOSN FETEVETESLTP POHYVDEEKLHEYYCQGFRTLVP
SKOTEAATYDVO FEKDALELLOSN FETEVENTLK ROCHALLAMST VETRE ESYDLALABHLTN
GCCK EGPOGLLVHRIGTTETAHREAKENGSCEFTGRLYPOFFOGFVAZEPTIKDA IL
SEETHPRIIGHTLEDD LEMPLEKVLFFMERAGVPLDVEFFA FLEALEFTELAVLTBETYDLS CRPFNIKSPYCLSDILYNE IDKAKSTPAEVLEALRSEHPIIEKLLS
TYVKALPKOVDSHTORIHP CAVYUKLACROPNLONIPIRGERGILLRKAFRLSEK
NSYFLSADYSQIELRFLAHLSQDKSLKFAFEGGEDIHAFTASOVFHVPLEOVSKEGRMQA
KTVHFGIVYGQQAFGLAKVLKISIGEXQELIQAYFSRYPELAHFVEETIOQAAKDLRVIL
MIGRERIIDSWNEFFGSRAAGGRFAVNTRIOTSAAELIKLAMLDISQAIKQOMKSRMLL
QIHDELLFEVPEEEIEEMORLVREKYESAMTLSVPIVVNILIGKMWAEC

704653 CPn 0613 SCHELLE TO ACCUMENTABLE ALLIA JUDON BIAMEZARA DAGGESKEDA KTAPIIAVIEMKOVIASSKNTAKTIONILEEGEKAPLKORVKOIVIIDMOLGGEVFEIDR IYSMLRFWKERKGFPIYIYVMOLCASCGYYVSCAATKIYATSSSLIGSIGVRSOFFINVK EGLNRYGVESOLLTAGKOKAPYNDYTPWTSHDREERQATLDFLYGQFVDIVTQNRPLLTK EKLVHTLGARIFSPEKAKQEØYIDVVGATKEQVLQDIVAVCKIEDNYRVIGSGGDGWWKR VASAAASSPLVTGMIKHDIYPLSHDAAYIPPYLAL

707435 CPn_0614 07435 705793
adt-ADP/ATP Translocase
VFIRHKVCKEFMOSSEV/PFSRLRAYLCPIYKSEFSKFVPLFLLAFFVGFNYCLLKNMKD
TLVIVGSDAGAEVIFF/KVWGIVPGAVIVTMYYGNLGSRYPROTUYYCFMAAFLGFFFLF
WIIIPYDODSLHLHSLGAKLGELLEPGGLRGFIYUMVRYWSYSIYYVMSELWSSVVLSNLFU
GLANQITTITEAGRFYALINTGLNLSSICAGEISYWMGKOTFVAYSFACDSWHSVMLNLT
MLITCSGLIMIWLYBRIHHLTIDTSIPPSRRVLAEEGAATANLKEKKFKRAKARNFLHL
IQSRYLLGLAIIVISYNLVIHLFEVVWKDQVSGIYSSHVEFNGYMSRITTLIGVVSVLAA
VLLTQQCIRKWGWYVGALVTPLVMLUVSGLLFFGTIFAAKRDISIFGGVLGMTPLALAAWT
GGMONVLSRCTKSTFDDYKEMAFIPLSPEDKHNGKAAI GDVYSRIGKSGGSLIYQGLU
IFSSVAASLANIALVLLIIMVVWIAVVAYIGKEYYSRAADAVATLKQPKEPSSSIVREAQ
ESVEQEEMAVL 705783

CPn_0615 / 708149 707634
pgsA-Glycerol-3-P Phosphatidyltransferase
LAXIMROFOKLLSLSRLWLALYFCQEKLHIRLLAIVGAMLSDVLDGYLARRYKATSRLGS
ILDPITDKYPFVCITVLYMEGSLSIAHLFFICARDLFLIIFVCYLSLVKGWKCYDYGSL
FWGKIFTVVOFIILLGVTAGGEIFWTGLVPLVALGFLYFLERIMDYKKOFLR

CPT_06/6 708704 710137
dnab-Replicative DNA Helicase
TLTTYYESSLIMDKSTGVPLPSPPHSKESEMIVLGCMLTGVHYLNLAANQLYEEDFYYLEH
KIIFKVLQDAFKQDKPIDVHLAGEELKRHNQITVIGGPSYLITLAEFAGTAAVLEEVVDI
IRSKSILRKMISTAKEIEKRALEOPKNVAEALDEAQNSFFKISOSTSVSQYTLVADKLRG
LTTYTDKPYLVQLQERGELFLQNAGGDNKSFFTGIPTHFIDLDQLIHGFSPSNLMILAAR
PAMGKTALAINIAENLCFQNRLPIGIFSLEMTVDQLIHRMICSRSSVDSKKISIGDLSGH
DGDRIVSVINEMGEHTLLIDDQDGLKVSDLRARARMKESYDIGFLIDVLQLLSGSGTL
RATESRQTEISEISRMLKTLARELNIPILCLSQLSRKVEDRANHRPMMSDLRESGSIEQD
SDLVMFLLRREYYDPNDKPGTAELIIAKNRHGSIGSVPLVFEKELARFRNYSAFECIS 708704 710137

710481 712316 CPT_U017 /10481 /12316
gidA-FAD-dependent oxidoreductase
LMWTHPIAYDVIVVGAGHAGCEAAYCSAKMGVSVLMLTSNLDTIAKLSCNPAVGGIGKGH
IVREIDALGGIMAEVTDOSGIGFRILNOTKGPAVRAPRAQVDKQLYHHMKRLLENTPGL
HIMQATVESLLDKEGVISGVTKEGMMFSGKTVUXLSSTFMRGLIHIGDRIFSGGRLDDP
SSOGLSEDLKKRGFPISRLKTGTPPRLLASSINFSCMEEQPGDLCVGFVHRTEPFOPPLP SSGLSEDLKRGF FISHLYGTFPRLASSINFSCMELQFGDLGVGFVHRTEPFGPELP QLSCFITHTMEKTRAIISANLHRSALYGGCIEGVGPRYCPSIEDKIVKFSDKERHHYFL PEGLHTQEIYANGLSTSMPFDVQYDMIRSVLGLENAIITRPAYAIEYDYIHGNVIHPTLE SKLIEGLFLCGQINGTTGYEEAAAQGLIAGINAVNKVFNRPPFIPSRQESYIGVMLDDLT TOILDEPYMFTGRAEHELLLRQDNACARLSHYGYELGLLSEERYELVKKONQLLEEEKV RLQKTFRQYGQSVVSLAKALSRPEVSYDMLREAFPNDIRDLGAVLNASLEMEIKYSGYID ROKILIOSLEKAESLLIPEDLDYKQITALSLEAQEKLAKFTPRTLGSASRISGIASADIQ VLMIALKKHAHH

712300 713010 CPT_0018 /12300 /13010

[pla-Lipoate-Protein Ligase A
KNMPTTNCIFLDLRGHSILHQLOIEEALLRVANQNFCIINSGAKDSIVLGISRNLNQDVH
ISRAQADHIPIIRRYSGGTVFIDSNTLMVSWIMNSSEASAQPQELLAWTYGIYSPLLPN
TFSIRENDYVLGHKKIGGNAQVIQRHRWHHTTFLWDIDLDKLSYYLPIPQQQPTYRNQR
SHEEFLTTLRPWFPSRDDFLERIKASGSLLFTWEEFLDNELEEILAQPHRKATTVLN

713462 CPH_0619 713462 /13013
ndk-Nucleoside-2-P Kinase
RRYVYTMEOTLSIIRPDSVSKAHIGEILSIFEQSGLRIAAMKMMHLSQTEAEGFYFVHRE
RPFFQELVDFMVSGPVVVLVLEGANAVSRNRELMGATNPAEAASGTIRAKFGESIGVNAV
HGSDTLENAAVEIAYFFSKIEVVNASKPLV

714145 CPn_0620 713519 CPT_0620 714145 713519
ruva-Holliday Junction Helicase
DKMYDYIRGTLTYVHTGAIVIECQGIGYHIAITERWAIECIRALHQDFLVFTHVIFRETE
HLLYGFHSREERECFRILISFSGIGPKLALAILNALPLKVLCSVVRSEDIRALASVSGIG
KKTAEKLMVELKQKLPDLLPLDSRVETSOTHTTSSCLEEGIQALAALGYSKIAAERMIAE
AIKDLPEGSSLTDILPIALKKNFSGVNKD

CPn_0621 . 714707 714144 ruvC-Crossover Junction Endonuclease LSRLGSSFKDNKFKVFQESIVSELIIGVDPOTIVAGYAIIAVEQRYQLRPYSYGAIRLSS DMPLPMRYKTLFEQLSGVLDDTQPNAMVLETQFVNKNPQSTMKLAMARGIVLLAAAQRDI LIFEYAPNVAKKAVVGKGHASKRQVQVMVSKILNVPEVLHPSNEDIADAFALAICHTHVA

CPD_0622 715761 714793
CT503 hypotherical protein
RYSVPELSTENEHERISESSESFHYYHOGREGMEHELLORWKDADIMEWOOICNILSGV
CORMOGRIM-LOKETODSTIQEHERIHLDYREGLIGALEEEYRREESAKHODLEKLQOENT
WLQHRIARKLQOIRHQADI LDEIKKELLOGVQRTEISBYRRCGYEHRIKKOLEEDLORYVS
OHDAHGI ELEEDKOIMAYARINDEKKOLIDLQEKDIYYKTYHBEIAKLDEKLOROGGAQ
TESEFGCITEKUTEVOTDLAEKKVATALLODIVEDOYCOLROLHKEKCMAMPONTKLOHLK
GLUSKEPESEVOVYESESKIZO

717011 CENTO-2013 ALTOHI ALTOHIO

CENTO-2014 PARTHER HER SAN TORKER HER TENNER HER TORKER HER T RELWEE: ORNE I VEOKKFOLLPPPAKL I SEVI GOTVVI I NALLGADDA I HFPETEEEPTSASFEESSAMFFPETS

ADLNESLQALVRESSDL

CPn_0624 718018 717011
gapA-Glyceraldehyde-3-P Dehyrogenase
AMKVYINGFGRIGRLVIRGILKRNSSVEVTAINDLVPGDALTYLFKFDSTHGRFPEDVRC
EADHLIVGKRKIGFLSERNVQNLPWKDLGVDLVIECTGLFTKKEDAEKHIQAGAKRVLIS
AMKGDICTFFVMGVNHKTENPEKDFVIGNASCTTNCLAPIAKVLLDNFGITEGLMTTVHA
ACKTRALOZICELEFFINNASCHERLDNI HAGDEN WAZFAVTLLDEELKFYLTTMAFREVELEDF PRVEDTYREEP JTTYDE CHAMKQAJETDEKYLEDYTDEQVVJJJOF IGGEYGS (FDALA) IALNDRFFKLVAWYDNETGYATRIVDLLEYVEKNSK

CPn_0625 718488 718060 rll7-Ll7 Ribosomal Protein wOMHARKFRVGRTSSHNRCMLANMLKSLIHYERIETTLPKAKELRRHADKMITLAKKNS LAARRIAIGRLMVRYNKLTSKEARQAKGGDTSVYNVDRLVVNKLFDELGNRFVERKGGYT RILKLONRIGDNAOKCIIEFLAS

CPn 0626 719670 718495 CPH_0626 719670 718495

rpoa-RNA Polymerase Alpha
WLPAKKKAQSVVLGKEKGMSDNAHNLLYDKFELPEAVKMLPVEGLPIDKHARFIAEPLER
GMGHTLGNALRRALLIGLEAPAIISFAMTGVLHEYMAIEGVIEDVTNIILALKGALLKKY
PMGDSSLGRTTQVLKASISIDASDLAANAGQKEVTLQDLLQEGDFEAVNPQVIFTVTQP
IQLEVVLRIAFGRGYTPSERIVLEDKGVYEIVLDAAFSPVTLVNYFVEDTRVGQDTDFDR
LVLIVETDGRVTPKEALAFSTQILTKHFSIFENMDEKKIVFEEAISIEKENKDDILHKLI
LGINEIELSVRSTNCLSNANIETIGELVIMPEPRILQFRNFGKKSLCEIKNKLKEMKLEL GMDLTQFGVGLDNVKEKMKWYAEKIRAKNTKG

CPn 0627 720059 719640 rsll-Sll Ribosomal Protein FLIRSRVLVKNQAQAKKSVKRKQLKNIPSGVVHVKATFNNTIVSITDPAGNVISWASAGK VGYSGSRKSSAFAATVAAQDAAKTAMNSGLKEVEVCLKGTGAGRESAVRALISAGLVVSV IRDETPVPHNGCRPRKRRV

CPn_0628 720461 720063
rsi3-si3 Ribosomal Protein
DAYTILREAGRMPRIIGIDIPAKKLKISLTYIYGIGSARSDEIIKKLKLDPEARASELT
EEEVGRLNSLLQSEYTVEGDLRRRVQSDIKRLIAIHSYRGQRHRLSLPVRGQRTKTNSRT RKCKRKTVACKKK

CP 0629 721881 720487 ser-Translocase KIREFRPYMTTLROFFLITELROKLFYTFALLTACRVGVFIPVPGINGELAVAYFKOLLG KIEFRPYMTTLROFFLITELROKLFYTFALLTACRVGYFIPVGINGELAVAYFKOLLG
SGÖLFOLDAIFSGGAFAOMTVIALGVVPYISASIIVOLFLUPPHALOREMRESSDOKR
REGRLTRLFTVALAVIOSLLFAKFALRMILTIFGIVLPTLLSSKLFGVPMIFYITTVVVM
TTGTLLLMMIGEGISDKGIGNGISLIIALGILSSFPSVLGSIVNKLNLGSODSSDLGLIS
IEEEALVFVFVLITTILIIEGVRKIPVQYARRVIGRREVPGGGSYLPLKVNYAGVIPVIF
AŞŞELMFPATIGGFIASESSMKKRIAALLAPGSLVYSICYVLLIIFFTYFTMTATOFHED
IASEMKKNNAFIPGIROGKPTOHYLEYTMNRVTLLGALFLAAIAILPSLLGCLLRVDSNV
SYELGGTAMLIVVGVVLDTMKQVDAFLLMRRYDSVLKTDRTKGRH

CEG 0630 722316 721885
rFIS-L15 Ribosomal Protein
MIKEESLFDISERKRRKKLLGRGPSSGHGKTSGRGHKGDGSRSGYKRRFGYEGGGVPLYR
RVPTRGFSHKRFDKCVEETTTGHAELFQEGEAITLDALKAKKAIARQAVRVKVILKGDL **EXTFVWODTAVVLSOGVONLLGIT**

CEn_0631 722812 722312 TESTS RIBOSOMA1 PROTEIN
TESTS RIBOSOMA1 PROTEIN
ERMSLSKNSHKEROLEEKVIVVARCSKVVKGGRKFSFSALILVGDCKGRLGYGFAKAMEL
TRAJARKGEAAKKNILAKIFALEROSIPHEVILVHHDGAQLILKPAKPGTGIVAGSRIVLIL
EMGIKDIVAKSFGSNNPMNQVKAAFKALTGLSPRKDLIRGAAIND

CPn 0632 723354 722827 CPI_0632 723354 722827
rlie-L18 Ribosomal Protein
KGTISSKULVNLLQVFAPNVLLNLIKVREFVMKMNMSVVKLVKLRKQAKNRSRVMESSLCK
KSHKRRRALRVRKVLKGSPTKPRLSVVKTNKHIYVQLIDDSIGKTLASVSYLSKLNKSQ
GLTKKNQEVAKVLGTQIAELGKNLQLDRVVFDRGPFKYHGIVSMVADGARXGGLQF

723760 723209 CPH_0633 /23/00 /23/09 r16-L6 Ribosomal Protein SMSRKAREPILLPCGVEVSIODDKIIVKGPKGSLTQKSVKEVEITLKONSIFVHAAPHVV DRPSCMGGLYWALISMWVGVHLGFEKRLEMIGVGFRASVQGAFLDLSIGVSHPTKIPIP STLQVSVEKNTLISVKGLDKQLVGEFAASIRAKRPPEPYKGKGIRYENEYVRRKAGKAAK

CPn_0634 724215 723787
rs8-S8 Ribosomal Protein
essikkrkiymontsosiadlitrirnalmaehlyvdvehskmreaivkilkhkgfvahy
lykeenrkamkyrloysdorkpviholkrvskpsrrvyvaakipyvfgnmgisvlsts
QGVmeGSLarskniggellclvw

724763 724206 CPIL VOSTOR TO TO THE TOTAL THE

CPn_0636 . 725100 72475 7124-024 Ribbsomal Protein PEREKEWAKKONIAYODAYFILAGNDKGKEGAVLSUTEDKVVVEGVNVRIKNIKRSQONPK OFFICIEAPIHISMVRLTTAGEPAKLSVKYTEQGRELMORRELDTSQLYRLVRGKKG

7114-1.14 Ribodomal Protein
TEIMIOUESOLKVADNTGAKKYKSEKYLOGGRRRYNTYGDVIVGSVRDVEENGGIKKGDV
PAN UVETRHITTEKINSSTLKEDTNYGVIIDDKGNEKSTRIEGEVAREIRDRGFIKISSL

Clm_06 08 725490 NKKEKVKSMASEPPCSRKVK VSAKMEKTYVRVER I FSHPQYLKVZRSSKKYYAHT ELKVSECDKVK I GETRPLSKLKRWRV I EHVCVÝS

725979 rl29-L29 Ribosomal Protein
ASCKGINMAAKKDLLTQLRGKSDDDLDAYVHENKKALFALRAENLLQNKVVKVHMFSTHK KNIARALTVKOEPYCKVHC

1116-L16 Ribosomal Protein
1116-L16 Ribosomal Protein
11MLMPKRTKFRKQOKGGFAGLSKQATFVDFGEYAMOTLERGWVTSRQIEACRVAINRYL
KRRGKVWIRIFPDKSVTKKPAETRAGKGKGAPDHWVAVVRPGRILFEVANVSKEDAQDAL
RRAAAKLGIKTRFVKRVERV

CPn_0641 727092 726409
rs3-s3 Ribosomal Proyein
kGRRIMGOKCCPIGFRTGVYMKWRSLWYGNKQEFGKFLIEDVRIRGFLRKKPSCQGAAGF
VVRMSGKIEVTIGTARPGLVIGKKGAEVDLLKEELRALTGKEVWLEIAEIKRPELNAKL
VADNIARQIERRYSFRRAMKKAMGSVMDAGAVGVXIQVSGRLAGAEIARSEWYKNGRVPL
HTLRADIDYATACAETTYGIIGIKVWINLGENSSSTTPNNPAAPSAAA.

CPn_0642 727440 727096 r122-L22 Ribosomal Protein RRHSMFRATARYIRVOPRKARLAGIARNLSVOEAEEQLGFSQLKAGRCLKKVLNSAVAN AELHENIKRENLSVTEVRVDAGPVYKRSKSKSRGGRSPILKRTSHLTVIVGEKER

CPn_0643 727725 727450
rs19-S19 Ribosomal Protein
EIRIMGRSLBKGPFVDHHLLKKVRAMNIEEKKTPIKTWSRRSMITPEMIGHTFEVHNGKK
FLTVFVSETAVGHKLGEFSPTRIFKSHPVKKG

728594 CPI_U648 /28594 /22722
riloasomal Protein
FIREINSMFKKFKPVTPGTRQLVLPAFDELTTRGELRGTKSKRSLRPNKKLSFFKKSSG
RDNLGHISCRHRGGGARQLYRVVDFKRNKDGITAKVVTVEYDPNRSAYIALLSYEDGEKR
YILAPKGIQRGDVVVSGESFFKPGCCTHILKSIPLGLSVNHIEDRFSSGGKLVRSAGLA
QVIAKSPGYVTLKMPSGEFFRHLNEGCRATIGEVSNADHNLRVDGKAGRRRWMGVRPTVRG

728933

TayINPVDHPHGGGEGRHNGYIPRTPWGKVTKGLKTRDKNKSNKWIVKDRRK

729636 728950 CFI_U0040
T14-L4 Ribosomal Protein
YREDLMVLLSKFDFSGNKIGEVEVADSLFADEGDGLQLIKDYIVAIRANKRQWSACTRNR
SEVSHSTKKPFKQKGTCNARQGCLASPQFRGGGIVFGPKPKFNQHVRINRKERKAAIRLL
LAQKIQTNKLTVVDDTVFVDALTAPKTQSALRFLKDCNVECRSILFIDHLDHVEKNENLR LSLRNLTAVKGFVYGININGYDLASAHNIVISKKALQELVERLVSETKD

CPn 0647 730490 729657 CFn_064/ /30490 /2965/ rl3-L3 Ribosomal Protein YLEYFSYCKNLPPLITCPFIFLRENFLFFLENSISKILSRFVSLFLQEESKSLLLMDKFM RSHISVMGKKEGMIHIFDKDGSLVACSVIRVEPNVVTQIKTKESDGYFSLQIGAEEMNAP AHTITKRYSKPKLGHLRYAGGRVFRFLKEVRGSEEALNGVSLGDAFGLEVYEDVSSVDVR GISKGKGFQGVMKKFGFRGGPGSHGSGFHRHAGSIGMRSTPGRCFPGSKRPSHMGAENVT VKNLEVIKVDLEKKVLLVKGAIPGARGSIVIVKHSSRT

CPn_0648 731636 730605
CT529 hypothetical protein
FFFKKPCKEVKNATNAIRSAGSAASKMLLPVAKEPAAVSSFAQKGIYCIQQFFTNPGNKL
AKFVGATKSLDKCFKLSKAVSDCVVGSLEEAGCTGDALTSARNAQGMLKTTREVVALANV
LNGAVPSIVNSTQRCYQYTRQAFELGSKTKERKTFGEYSKMLLTRGDYLLAASREACTAV
GATTYSATFGVLRPLMLINKLTAKPFLDKATVGNFGTAVAGIMTINHMAGVAGAVGGIAL
EQKLFKRAKESLYNERCALENQQSQLSGDVILSAERALRKEHVATLKRNVLTLLEKALEL
VVDGVKLIPLPITVACSAAISGALTAASAGIGLYSIWQKTKSGK

732672 731710 CPT_U649
fmt_Methionyl tRNA formyltransferase
LNLKVVYFGTPTFAATVLQDLLHHKIQITAVVTRVDKPQKRSAQLIPSPVKTIALTHGLP
LLQPSKASDPGFIEELRAFNADVFIVVAYGAILRQIVLDIPRYGCYNLHAGLLPAYRGAA
PIQRCIMEGATESGNTVIRMDAGNDTGDMANITRVPIGPDMTSGELADALASQGAEVLIK FIGHT INBUALESCALEVITY AND AGRICULTURE OF THE PROPERTY OF THE

CPn_0650 733513 732665

lpxA-Acyl-Carrier UDP-GlcNac O-Acyltransferase
SRRMMASIHPTALIEPGAKIGKDVVIEPYVVIKATVTLCDNVVVKSYAYIDGNTTIGKGT
TIWPSAMIGNKPQDLKYQGEKTYVTIGENCEIREFALITSSTFEGTTVSIGNNCLIMPWA
HVAHNCTIGNNVVLSNHAQLAGHVQVGDYAILGGMVGVMGVMGVRIGAHAMVGALSGIRRDV
PPYTIGSGNPYQLAĞINKVGLQRRQVPFATRLALIKAFKKIYRADGCFFESLEETLEEYG
DIPEVKNFIEFCQSPSKRGIERSIDKQALEEESADKEGY/LIES

733975 CPn 0651 733517 LidZ-Myristoyl-Acyl Carrier Dehydratase
MNQBOVIKHRELLDLLPHRYPFLLVDKVLSYDIEARSITAQKNVTINEPFFMGHFPNAPI
MFGVLIELALAQAAJVLIGULEIDRINKRIALFLGIQKAKFRQAVRFGDVLTLQADFSLI
SSKCKKAWAQARVDSQLVTEAELSFALVDKESI

CPn 9652 714880 CPI_052 7 NA80 733990

LDXC MYLISTOY LOLONG DEGGET/LSS

KROLLYODSLAGYMERTORTEKREVRYSSIVG HIDJKOSTLHLOPACTNY/LVFOROS

ACGITERVPALLIPHYYTTHETTELSROSAV LATVEHLMAALRSINI DINLI LOCGGER PLI

GOGGONYVELLIQVAG LOEGEDRYG LARLTPPYYYOHODIFLAAFFODELK LOYTLHYPO

COTTOTOYKOLV INEEDFROELALCRIFFALYNELOFLMEKGLIGGGCLDNAVYFKOTGII

GRODGRADEPVRIK LLOLIGDISCHGRPFVAHVLAVGGGHSSNIAFGKKILEALEL

CPrs 9653 736439 734868

curE-Apolipoprotein N-Acetyltransferd
GEPVLRIFCFVISMCLIAFAQPDLSGFVSILGAACGY
SLEPLKKPSLPLRTLFVS
CFFWIFTIEGIHFSWMLSDQYIGKLIYLVWLTLITILSVLFSGFSCLLVALVRQKRTAFL CFFMIFTLEGIHFGMMLSDQYIGKLIYLVMLTLITILSVEFSGFSCLLVAIVRQKRTAFL
WSLPGVWVAIEMLRFYGIFSGMSFDYLGWPMTASAYGRQFGGFLGWAQQSFAVIAVNMSF
YCLLLKKPHAKMLWVLTLLLPYTFGAIHYEYLKHAFQQDKRALRVAVVQPAHPPIRRKLK
SPIVVWEQLLQLVGPIQQPIDLLIFPEVVVPFGKHRQVYPYESCAHLLSSFAPLPEGKAF
LGNBDCATALSQHFQCFVI IGLERWYKKENVLYWYNSAEVISHKGISVGYDKRILVFGGE
VILMOKFGGLIGPDIFPEYALAGKRUPGREGATAYNGRAPRIGITICYEETFGYRLOSGE
VILMOKFGGLIGPDIFPEYALAGKRUPGHIGMLPEYJEFSMM"VRAGOTGYTAAVDGLOFILK
LGLYBUTHETMIZWYJEFSHARVIFFLIGMLPEYJEFSMM"VRAGOTGYTAAVDGLOFILK
LLYBUTHETKARDOVLETGLEFMYKTLYGYGYGYTFNILLAFGAVOYYDXXFLGYRLLAK

CPn_0654 737051 736503
vdlD/ycia-acyl-Coa Thioesterase
KKIIDFLSVDRYYRNQEYPIKILSVESTMLKKKPVSFSCIDGHIYKIFPNDLNANNTVFG
GLLMSLLDRLALVVAERHTESVCVTAFVDALBFYAPAYMGENLICKAAVNRTWRTSLEVG
VKWAENIYKQERRHITSAYFTFVAVNEDNQPIPVHQIVPETPEEKRRYNEADRRRQARL

737856 737101

CPT_0655
73/856
73/101
dnaQ-DNA POl III Epsilon Chain
KEIMSLLKDTVFTCLDCEMTGLDVKKDRIIEIAAVRFTFDSVISSIEFLINPERVVSAES
QRVHHISNAMLRDQPKIAEVFPOIKAFFKEGDYIVGHSVGFDLQVLAQEMERIGETFLSK
YTIIDTLRLAKEYGDSPNNSLESLAVHFNVPYDGNHRAMKDVEININIFKHLCKRFRTLE OLKOVLAKPIKMKYMPLGKHKGRCFSEIPLAYLOWASKMDFDSDLLFSIRHEIKHROKGT

CPn_0656 737842 738048
No robust homolog present in Genebank/EMBL as of 11/7/98
THNFLLLPLSLFDILLTVEGFLCLTLYFASVQRMPCEQKRVPGNLYYYYIAAHSSLCLSV

738476 738051

YjeE (ATPase or Kinase)
PMGRYRRVSHSSQETLLLGTELGQVLVPGAVLLLFGDYGAGKTEFVRGIVSGYLGDTIAE
EVASPSFSILHVYGNEPKRLCHYDLYRIDQKNQEYIFQDAEEDDVLCIEWADRLPKPRFC DTINIYITMOTNMEREIIIEKR

CPn_0658 739180 738455
CT538 hypothetical protein
KRYGMDISGAVKQKLLQFLGKQKKPELLATYLFYLEQALSLRPVVFVRDKIIFKTPEDAV
RÄLEDDKKIWRETEIQISSEKPQVNENTKRIYICPFTGKVFADNVYANPQDAIYDWLSSC
POMMEKQGGVRIKRFLVSEDPDVIKEYAVPPKEPIIKTVFASAITGKLFHSLPPLLEDFI
SSÄLRPMTLEEVQNQTKFQLESSFLSLLQDALVEDKIAAFIESLADDTAFHVYISQWVDT EE

CPA-0659 739482 739838
LEXA-Thioredoxin
LOEDRADSNSIFREGKLMVKIISSENFDSFIASGLVLVDFFAEWCGPCRMLTPILENLAA
ELPHVTIGKINIDENSKPAETYEVSSIPTLILFKDGNEVARVVGLKDKEFLTNLINKHA

740327 739860

Specification with lase which control the control of the control o

741139 740327 C###20001

#ip-fKBP-type peptidyl-prolyl cis-trans isomerase
##\$P#ELKIKDRRKMNRRWMLVLATVALALSVASCDVRSKDKDKDGSSLVEYKDNKDTNDI
ELSDNOKLSRTFGHLLARQLRKSEDMFFDIAEVAKGLQAELVCKSAPLTETEYEEKMAEV
QK##FEKKSKENLSLAEKFLKENSKNAGVVEVQPSKLQYKIIKEGAGKAISGKPSALLHY

KGSFINGOVFSSSEGNNEPILLPLCOTIPGFALGMQGMKEGETRVLYIHPDLAYGTAGO
PFNSLLIFEINLIQASADEVAAVPQEGNQGE

742938

CPHE 0662 742938 741172

asps_Aspartyl trna Synthetase
skGeymkyrthroneltsnhigenvolagwyrryrnhggvyfidlrdrfgitqivcrede
QPELHQRLDAVRSEWVLSVRGKVCPRLAGMENPNLATGHIEVEVASFEVLSKSON/FFSI
ADDHINVNEELRLEXRYLDMRRGDIIEKLLCRHQVMLACRNFNDAQGFTEIVTPVLGKST
PEGAABYLVPSRIYPGKFYALPOSPOLFKOLLMVGGLDRYFQIATCFRDEDLRADROPEF
AQIDIEMSFGDTODLLPIIEQLVATLFATOGIEIPLPLAKMTYQEAKDSYGTDKPDLRFD
LKLKDCRDYAKRSSFSIFLDQLAHGGTIKGFCVPGGATMSRKQLDGYTEFVKBYGAMGLV
WIKNQEGKVASNIAKFMDEEVFHELFAYFDAKDQDILLLIAAPESVANQSLDHLRRLIAK
ERELYSDNQYNFVWITDFPLFSLEDGKIVAEHHPFTAPLEEDIPLLETDPLAVRSSSYDL
VLNGYEIASGSQRIHNPDLOSOIFTILKISPESIQEKFGFFIKALSFGTPBHLGIALGLD
RLVMVLTAAESIREVIAFPKTQKASDLMMNAPSEIMSSQLKELSIKVAF

CPn_0663 744220 742901
his3-Histidyl trna Synthetase
KSNHFERRHHYTVTLPKGVPDIFPYLADAKQLWRHTSLWHSVEKAIHTVCMLYGFCEIRT
PIFEKSEVFLHVGEESDVVKKEVYSFLDRKGRSMTLRPEGTAAVVRSFLEHGASHRSDNK
FYYTLPMFRYERQOAGRYRQHHOFGVEAIGVRHPLRDAEVLALLWDFYSRVGLOHMOIQL
NFLOGSETRFRYDKVLRAYLKESMGELSALSOGRFSTNVLRILDSKEPEDGEIIRQAPPI
LDYV3DEDLKYFNEILDALRVLEIPYAINPRLVRGLDYYSDLVFSATTFTQEVSYALGG
GRYDGLISAFGASLPAGGFGVGLERAIQTLLAQKRIEPOFPHKLRLIPMEPDADOFCLE
WOOHLRRIGIFPEVDWSHKKVKGALKAASTEOVSFVCLIGEREVISQQLVIKNMSLRKEF

CIT_0664 744775 744557
NO cobust homolog present in Genebank/EMBL as of 11/7/98
LWEAHAMKKLIALIGIFLYPIKGNTNKEHDAHATVLKAARAKYNLFFVQDVFPVHEVIEP

CTD_DIGS 744998 746365

OBJAN HEROSPHOEDBLIFF TEGROPOTT

EMBAWTKEFOPEKHEKELEDQEVVKKKYKYWRERIFFOMFIGYLEYYFTRKSFTFAMPTL

LABIGFDRAQUALIGOTOFFYGEIGKEYGOWADGUARDWARGEGMETGEMETGEMETFOMSS

ELYLFALIWW BLANGWEGARGWEFGARGELITHWYARGEFOTHWEIWETGHINGGALFFILDGF

ELDYGANGGAMYVISTELCIGMGLVLINRERDTFQ/IZJEPFIEKYKRDPHHAHHEGKSAGE

GTEELERELETRELETYVLYNOWLWFLAAAGFFIZIVMAVNDWGALFLIETKHYAAVK

KISKONEGEMENVLEGLÖLLEFAILGMWEGRSHNOWWV ANFCVSLFEIGGLFGMLVA KISKGNEGEMNVI DGTLLFVIGFFLYGPOMMIG BAELSHKKAAGTASG WGWKGFFIALLACASIALLLFLFTWNAJEKNTRSKA

746379 CPn 0666

CPn_0666
dnat-DNA POI III Alpha
GFFLTWIPLHCHSOYSVLDAMSSIKOFVAKCQEFGIPALALTDHGNLYGAVDFYKECTQK
GIOPIIGCECYIAPGSPFDVKKEKSSRAAHHLILLCKNECGYPNLCILTSLAFTEGFYYF
PRIDNDLLPY/GERLUGLEVILLYCLEVALE GRADE GERALLELLWEGULFYLOGTEVZLIK
MEREIJAGFKEEWLKGEYULLIYA GRADE SYVEKSBAOMAELFKDIPEVINNILDWQAH MCHESTAGE AND ANALYSM AND ANALYSM AND ANALYSM AND AND AND ANALYSM AND AND ANALYSM AND ANAL

CPn_0667 751097 750177
No robust /nomolog present in Genebank/EMBL as of 11/7/98
NISLLCKIOKRYFMKKLILYFAAFVASLFCGVFLWDRVPCAQKIMRLAADHSSEVFSKSC
RFVRKISGFEELQVFERWYSEQALALFPEYRDGKSFVELAFIPHTLMMVRFSKEEPVKK
HIISQEGFILWSLVMGEWYLHTGTWTCSKGFRECLLLHAGKQDMRVIQTLATLGGTTSRE
SLAQALALKNIRAERVIKECQKKKLIFASGNQIGTHFQOFQPIRGCTTTLINNPVWLQKP
RHAAVFPAQYSEDRVRHLVKMIFGDNFLIVRSSMVYVPVYKISLVSADNSVRVEYINAVT GKSFORL

CPT_0668 751176 752162
CT547 hypothetical protein
wRFWVVSPRLIMKFLLYVPLLLVLVSTGCDAKPVSFEPFSGKLSTGRFEPGHSAEEYFSQ
GOFFLKKGNFRKALLGFGIITHHFPRDILRNQAQVLIGVCYFTQDHPDLADKAFASYLQL
PDAEYSEELFGHKYAIAQRFAQGKRKRICRLEGFPKLANADEDALRIYDEILTAFPSKDL gaqalyskaallivkndlteatktlkkltlofplhilsseafvrlseiyloqakkephnl Dylhfaklneeamkkohpnhplnevvsanvgamrehyarglyatgrfyekkkkaeaaniy

752140 752775

YRTAITNY POTLLVAKCOKRLORISKHTS

Pn_0670 752738 753196

rsbW-sigma regulatory factor-histidine kinase PRRLLNRYTMTFFEGETVFPAVLSELHSMLDLIKRAGKQSKCPQEKLLKLELACEELLVN IISYAYQGENSPOTIAISCISHRGDLEVVIKDHGPSFNPLAVSINIQEDLPLEQRKLGGL GIFLAKSSVDEFLYAREDHCNIVHLKMLNGQHS

CPn_0671 753660 753205
CT550 hypothetical protein
RITINORKYTMSLDFFEEFYHOSILNTGTSFPEGYLNIAEILSYPHCTDANTDFLCSQSD
MDFIIAESKDKLTLFNADFAIWLVPELVQGQAVTRGYIAVSQGEGNYEPEMAFEASGQYN
QSSLILEALQLYLKDIKDTENALRSFRFNNDH

CPn_0672 753723 755048
dacF(pbp5)-D-Ala-D-Ala Caroxypeptidase
TIKSPHMKRPFFTYLCIIFYGSCASLSLHAGLSFPEVRGATAAVVHADSGKVFYDKDIDA
VIYPASMTKIATALFILKHYPTVLDTLIKVKQDAIASITPQAKKQSGYRSPPHWLETDGS
TIQLHLREELLGWDLFHALLVCSANDAANVLAMACCGSVEKFMDKLNFFLKEEIGCTHTH
FNNPHGLHHPNHYTTTRDLISIMRCALKEPPFRGVISTTSYKIGATNLHGERILSPTNKL LLPGSTYHYPPALGKTGTTKTAGKNLIMAAEKNNRLLVTIATGYSGPVSDLYQDVIALC ETVFNEPLLRKELVPPSDCLQLEIANLGKLSCPLPEGLYYDFYASEDREPLSVSFIAHAD AFPIEQGDLLGHWVFYDDEGKKISSQPFYAPCRFERTIKPWKLYMKRVFTSYRTYMSITM

LLMYFRIRKHRKYKNLKHYSKI

CPn_0673 755242 755463
CT552 hypothetical protein
GKSTECKAYHCFLKQVSIALNREEVWDNPHHLMFILMQFQQFSGEQDRFGSFLEATIRDR
VSFLVLQEKIATLK

756683 755577

CPn_0674 756689 755577
fmu-RNA Methyltransferase
RGILYVTMVPFRQHHAYQLLKQLHTSAISEADRVSYYFKQNRSLGSKDROWIQNIIFNIL
RHRRLLETLILDSGEQVTPEALVAKVNEGVLENLDSYSAIPWEVRYSISDDLAHFSKHHPUQST
GEEQABEIAKIMLTEATTIRWTTDKISVKELQEKLEYPSSPGELPEALHFSKHHPUQST
EAFRRGFFEIQDENSQRISQJICHTDKDIVLDFCAGAGGKSLIFAQKAKHVVINDSRKAI
LQTAKHRLLRAGARNFSLAPQLRIDSFCVVIVDAPCSGTGVFRRHPEHKWQFSKKLLLNY
VRVUKSILKQASAYVGPRGRLVYITCSLLKEENEAHVAYMHSLWKEVHRKTLPLQVGKG
DAFFTSHFCKI

757931 756768

CPR_0675 75/931 756768
CT696 hypothetical profess
VPLOMILDFORSIGYYLRVLELAIRDFPRILAYDRKRULLDAWPYNDFUPTNYDTSVSTI
RQVIHELFSWSAILYSISSRILAIIFELFURKERSWSDFFHEE/SWSHIKKAIVORKUM
RKGULLEESKRPVKKIVQAANKYEGKORONESSWSDFFHEE/TVSEVAFFLAIGEVGRELAA
DAGUMIIEALTTULEGHTAYUFUSTLERUNQFIGEKAQFIKTLSEKGYVULRELIQUFSU
GAEDFQTIIMSIISDSUSEVLAIGUIFERSKTFWILWQETALASHEDSKKAGFL
AEVLRKVIVEKKUHVGKSPNTTPEE/FWIIYGIIRDJNPALWDKMITMILMRWILDYDRDIG
IALRKAAEYYNPHPSFWRQFIRLWOPPP

CPn_0677 760410 759256
No robust homolog present in Genebank/EMBL as of 11/7/98
RIAMGINPSONRSPDDWWRGAGGDSSSTOGTGATNSNLGAHNVTTSTSOPQVASKAKQL
WQTVREFFLGKKSPDSSQGASGPAMQSPSGPTIEPTRPAPPPPTTGGANAKPATHCKGR
APQPPTAGSSSGSEQPTAMSSEVAKLVSELKDAVHSHAESQKVLKKVSQELQTKMTDMEN
NRCPDYLLHGYRVIARALQOTYTEQSMLIEGTSSTGPVPQAVTVAKDAVTQTVRGATAKL
ENPKPGNDPDGVLMQVVISLGIEGFLDFGESIGNFLETRYSDFGGDDSDIDYTSDIANL
ESALDRVREMHENEMPRIMIALARELGAAVHSHATSVRIANAGKNHTDVVRMANESSRL
LQCMKVLSVGAWANTMTVLIGDLFE

CPn_0678 761329 760682
No robust homolog present in Genebank/EMBL as of 11/7/98
KIIMSVNPSGNSKNDLWITGAHDQHPDVKESGVTSANLGSHRVTASGGRQGLLARIKEAV
TGFFSKNSFFRSGAPRSGQOFSAPSADTVRSPLPGGDARATEGAGRNLIKKGYQFGMKVT
IPOVPGGGAQRSSGSTTLKFTRPAPPPPKTGTTNAKRPATHGKGPAPQPPKTGGTNAKRA ATHGKGPAPQPPKGILKQPGQSGTSGKKRVSWSDED

CPn_0679 762936 761725

pgk-Phosphoglycerate Kinase
GYMDKLTVQDLSPEEKKVLVRVDFNVPMQDGKILDDIRIRSAMPTINYLLKKHAAVILMS
HLGRFKGGGFQEEYSLQPVVDVLEGYLGHHVPLAPDCVGEVARQAVQLSPGRVLLLENL
RFHIGEEHPEKDPTFAAELSSYGDFYVNDAFGTSHRKHASVYVPQAFGRAAAGLLMEK
ELEFLGRHLLTSPKRPFTAILGGAKISSKIGVIEALLNQVDYLLLAGGMGFTFLQALGKS
LGNSLVEKSALDLARNVLKIAKSRNVTTVLPSDVKAAENLQSKEYSVISIDQGIPPHLQG
FDIGPRTTEEFIRINGSATVFMNGPVGVYEVPPFDSGSIAIANALGMHPSAVTVVGGGD
AAAVVALLAGCSTKVSHVSTGGTASLEFIFGGGTLRTTEVLSPEKS AAAVVALAGCSTKVSHVSTGGGASLEFLEQGFLPGTEVLSPSKS

764254 762971

CPT_0680
764254
7904-Phosphate Permease
YSMLPLIIFVLLCGFYTSMNIGANDVANAVGPSVGSGVLTLRQAVVIAAIFEFFGALLLG
DRYAGTIESSIVSVTNPMIASGDYMYGMTAALLATGVWLQLASFFGWPVSTTHSIVGAVI
GFEENLOKKGTIIVMNSVGIILISWILSPFMGGCVAYLIFSFIRRHIFYKNDPVLAMVRVA
PFEAALVIMTLGTVMISGGVILKVSSTPMAVSGVLVCGLLSVIITFYVHTKHCSYISDT
PKRGSLTYRLKERGGNYGRKYLVVERIFAYLQIIVACFMAFAHGSNDVANAIAPVAGVLR
GAFFASYTSYTLIRLMAFGGIGLVIGLAIWGMRVIETVGCKITELTPSRGFSVGMGSALT
LAZESILGLPISTTHVVVGAVLGIGLARGIRAINLNIIKDIVLSWFITLPAGALLSILFF FALFALFH

CHI 0681 765001 764258
CT691 hypothetical protein
NG#KSHKSPTRSFRQVIIAKKAILMOTLARLFGQSPFAPLQAHLEMVVSCVEYMLPIFTA
REDGRYEELLEMAKLVSDKEYQADCIKNIDMRNHLPAGLFMPISRAGILEIISIQDSIADT
AEDANAILLHIRRLNFYPSNETLFFFFLEKNLEAFELTMTLLHEFNQLLESSFGGRKADKA
RLIAGRVAKSEHESDVLQRELMQIFFSDDFIIPEKEFYLWLQVIRRTAGISDSSEKLAHR INMTLEEK

764912 765955 CPT. 0682 764912 765955

DDD-ABC ATPASE DIPEPTIGE TRANSPORT

TSKCLHKNISLFRNNNLPKRSCKRLMASNPILQIEDLSITLAKQRQQYPIVQSLSFTINEG

OTLAIGESGSGKSVSAHAILRLLPCPPFSVSGQVMFQGHNLLTASRSIQKKIIGTEISM

IFONPQASLMPVFTIEQGFREIIHTHLALTAEVAKEKHLYALEETGFHDPRLCLNLYPHQ

LSBGMLQRICIAMALLCSPKLLTABEPTTALDVSVQYQTIQLLKTLOKKTGMSLLIITM

MGVVAETADDVLVLYAGRMVECAPAVQMFHNPSHPYTRDLLASRPSLQPQQLGSFNPFG

QPPHYTAFPSGCRYHPRCSKILNRCSAEAPETYPVREGHKVRCWLYDD CPh 0682

765936 766919 CPHE_D683 765936 766919
dppR_ABC ATPase Dipeptide Transport
GVGGMTINFPQPLIQATSLTKHYYKRSFWFQGKTIASRPVDDVSFSLYSRRAVGLIGESG
SGKSTLALALAGGLIPLTSGFLIFFNOTPIKLHSKHGGHQLRSQVRLVFONPQASIAPRKTI
LDSLGHSLLYHKLVPKEKVLATVREYLELVGLSEEYFYRYPHQLSGGQQQRYSIRRALIG
VPQLIICDEIVSALDLSIQAQILNMLAELQKKLSLTYLFISHDLAVVRSFCTEVFIMYKG
GIVEKGNTKRIFSDPQHPYTRMLLNAQLPETPDQRQSKPIFQEYHKDSEESSTGCYFYN RCPOKOEACKSEIIPNOGDAHHTYRCIH

768056 767181 CPD_9584 /68056 /67181
spoJ/parB-Chromosome Partitioning Protein
EKSGDIVTEEISKDTIIEVAIDDIRVSPFQPRRVFSNEELQELIASIKAVGLIHPPVVRE
ICTGDRVLYYELIAGERRWANQLAGATTIPVILKHVIADOTAABATLIENIQRVNLNPI
EMAEAFKRLIHVFGCITODKVAYKVGKKRSTVANYLRILALSKTIQE\$LLQQITLGHAXV
ILTLEDPILREKLNEIIIQEHLAVREAELIAKQLISEEGSSIELKFTPLDMAESSKQHEE
LQQRLSDLCGYKVQIKTRGSKATVSFHLQNTQDLQKLEAWLSSHGTLSESLS

CPn_0685 768016 768217

No robust homolog present in Genebank/EMBL as of 11/7/98

FPOSQYLLIFPNRILDLQAFEILDVQGMLTDQRKHIQMLHKHMSIEIFLSNMVVEVKLFF CPn 0685

768373 768176 CPn 0636 tra_0000 No robust homolog present in Genebank/EMBL as of 11/7/98 AKDSMMPGGRLFRV7QELFFFSSVYVCEQRRPRKLYPSLOHLNFPIEKPRFLLKGFKKEL

768501 769214 CPE_9687 768501 769214

TT4B2 hypothetical protein
RK HKRILMIAYRESTENGESEMCKLUHNIWEKEYSESSATATCIVLASELSEK IVSNTYK
HEGAKENSI ELLUTRAAEVAVSGGELPSKSALSSLEGAYHLGESMKPYAGELASEFY THN
EPLIGAYYA:LAYREGALQLEHFIGKLIKETSEAGADQLYDVALSKSYQLLQTANSSPE
YPTLSELTILIKVIELKELLHODVSQDFAALKSSELFHOFERMYSDGEWTLSKREGKKG

760376 CIm. 0688 CT481 hypotherical profein SIMLIVLAFROYFFGHSRS NYLBULKONFAITLEKERTIKGHGHMLTFGFASFD FYTNIFPFLEEQKIPAVOV TOSNAQDLHGHRUKGGFTLAFGDEIFGNYMPFCC ONELIEMAKSPYIQLASSGFAIRNLMNYPPYLTTEILUSRHHIETITGKPLAFLFPFGK SDPSKKLAADHYPYSFLLGNTINKKIKTHNIYRLDIKPMQYVCPSLFQSSRYLKNWIKE KSKQLYLKKQLPKR

770147 771407

CPn_0699 771407 / 770147
yfho-nits-related Aminor/Finsferise
white year of the complete of the

771436 h72704 CPn_0690 /772704 771436
ABC Transporter Membrane Protein
LSVLRGDKVLVSIETF\$SIASGSPVOKAAEACYTQYSKOPSSKEVLSSFSWIQELSLFPD
RYNLATGASELIKOHYLHINNISLAFECILINGKYEPSELSQLPBGVIVCGIDEARGSLSSF
WGGFDVNKHPLAFLAVCSEDGRVIYIFDEEMQTSDPIFVRHISFPTVSDHDVIFSPRIV
VILOQRASAQIQISADVDLEHWGSSKTIVNGVTELFVGEGADLTVTMVPGYSEEDTLSWS
TIATVEKDAICRMYONLLESCOGFGWFDNTSYIVGKKGHAESLVLVQSPRKTWANNLMS
DAEETVSRONIKSILYSCHFLFECTISISSCODLSDANGKHDTLLLSSEARVSTFPRLEI
ETDEVKASHGATVGPLDPOQIFYMRSRGMTEAEAQEKLIHGFLKQGLVSDTFLGSSFQLN

772685 773467 CPn 0691 CPT_0691 //3467 //2003 CT691 hypothetical protein RGLGSMLK/KHLHASCNDVKILDDFNLNIOPGTMHVIMGPNGAGKSTLAKILAGDESVLV SSGEIALOGONLLSMLPEERSRAGLFVOFOMPPEIPGVNNKMFLRDAYNARRRANQEGDI SIDEFNTLISTVLETYENARTDLFLDRNVNEGFSGGERKRNEICOMLVLEPEMVLLDEP DSGLDVPALRLICRVLEKYRELHPTSSLCIVTHNPKLGNLIRPDVVHLLLDGRVALSGDV

SLMHELEAKSYOEVTKRVAWR

773461

774945

CPn_0692 ABC/Transporter ABC/Transporter
IQFCATGLKVMGESVKVFLEEREDYPYGFVTPIESQGLTRGLSEETIEEIAALRNEPQF
IQFFCATGLKVMGESVKVFLEEREDYPYGFVTPIESQGLTRGLSEETIEEIAALRNEPQF
IJFFRQAYRYMKQLHEPAWARLHYGPTAYDDIVYFSSPKQKKPLGRLEDADPEILDTFK
KI/SIPLDEOKRLLINVENVAVDLVFDSVSIGTTFKEALEKAGVIFCSLGEATQEHPNLVKK
V/LGSVVSHRONFFAALINAAVFSDGSFVTYPKGVKCPMDISTYFRINNEAQGFERTLIVV
EDGGYASYLEGCTAPAYSSNQLHAAVVELVAHEHAVIRYSTVQNWYAGDKTGKGGIVNF
VTKRGLCAGYRSKISMSQVEVQAAITMKYPSCILKGDESVGFYSVALTSGKMQADTGTK
MLHVGKRTTSTVISKGISSDESKNTFRSLVSLGKKAEHSSNYTQCDSMLIGKASGAYTDP
KIVVENSTSSIEHEATTSKLREDQLLYLRSRGLSPEEAVSLVIHGFCREIIEQLPLEFAQ EASKLLLIKLENSVG

775240 _0693 776292 CRT_0693 776292 775240

TPR Repeats (0-Linked GlCNAC Transferase homolog)
LKSTNHVLGEISMEERAKHLAKEFLCSGINLFLSGEYEQAEKRLKETLELDSTAALAYCY
LKGINAVLGEISMEERAKHLAKEFLCSGINLFLSGEYEQAEKRLKETLELDSTAALAYCY
LKGIIALETGRVSEALNWCSKGLASEPGDSYLRYCYGVALDRGNQYEAAIEQYSAVVALHP
DDVECWFSLGSVYHRLKRLQEALDCFDKILALDPWNPOSLYNKAVILSEMDDEAESIRLL
EVAVAKNPLYWKAWVKLGFLLSRSKRWDKATEAYERVVQLRPDLSDGHYNLGLCYLTLDK
TRLALKAFQEALFLNAEDADAHFYVGLAHLDLKQMREAYEAFNSALSINLEHERAHYLLG
YLHHMOGETDKATKELLFLQKKDSTFAPLLQKTVVSDPSSMQFERRLDTIS

CPn_0694 779635 776330

pbp2-pbp2-transglycolase/transpeptidase
FSDESEANNIHSMKPKKFPIYLSIAQKTNRLLSGIVIAFAVIALRLWYLAVVEHEQKLE
EAYKPQIRVLPQYVERATICDRFGKTLAVNQLQYDVSVAYGAIROLPTRAWRVDEHGHKQ
LIPVRKHYIMCLSELLSQELHLDREAIEDAIHAKASVLGSVPYLVAANVSERTYLKLKML
KKDWPGLHVEAVVRHYPQESVASDILGYVGPISLQFYKRYTYQELSQLRECVTAYEEGED
PKLPEGLASIDQVRALLESVESNAYSLNALVGKMGVEACWDSKLRGKIGKKPILVDRRGN
FIQEMEGAVPEAPGTKLQLTLSAELQAYADALLLEYEKTETFRSAKSLKKREKLPPLFPW
HKGGAIIALDPNNGEILAMASSPRYNNDFVNAKVAEDSKAVRSSIYRMLENKEHIAEIY
DRKYPLIRERNPLTGLCYEEILPLTFDCFLDFLFPENSVIKLQLKRNSFVQQAIEVQNL
VTRLLSLFPYEEGTCPCSAIFDAVFPNEEGHTLIQEVISLQEDKWHMECLNQHKADIEEL
KEALDQVTNELPANYDKILYTDILRLIVDPERFSPVLPSEVHRLSLSEFTELQGRYVVLR
SAFSTILEDAFIEVHFKSWRKSEFLQYLAAKRQEBALRKORYPTPYVDVLEEEKTRQYKM
FCQEHLDTFLAYLFSKTPYKEGLEPYYDILDLWINELDNGAHRALSWHEHYLFLKERVSK
ESHLPALFSTFREFNELQRELLGKYPISIVRNKROTEQDLAASFYPVVGYGYLRRHAYG
QAATLGSIFKLVSAYSVLSQRILWGHNEEPANPLVIIDKNSFGYRSSKPHVGFFKDCTPI
PTFFRGGSLPCNDFMGRGFIDLVSALEMSSNPFYFSLLVGEGLOPDELLADASSLFFFEDE
TGLGLPGEYAGRVPHDLAYNNSGLYATAIGQHTLVVTPLQTAVMLASLVNGGVVYVPKLL
LGEWEGEHVSYLSSKKKRTIFMPDAVVEVLKTGMRNVINGQVGTARAIGSGFPPQLLSRI
IGKTSTAESIHRVGLDREFGTMKMKDIWFAAVGFSDQDLSLPTIVVIVYLRLGEFGRDAA
PMAVKMIDMWEKIQQRESFLRG 776330 779635 PMAVKMIDMWEK IQQRESFLRG

780201 781382 CPn_0695 homologous to CT695 homologous to CT695
SLEVSMKKLLKSALLSAAFAGSVGSLQALPVGNPSDPSLLIDGTIWEGAAGDPCDPCATW
CDAISLRAGFYGDYVFDRILKVDAPKTFSMGAKPTGSAAANYTTAVDRPNPAYNKHLHDA
EWFTNAGFIALNIWDRFDVFCTLGASNGYIRGNSTAFNLVGLFGVKGTTVNANELPNVSL
SNGVVELYTDTSFSWSVGARGALWECGCATLGAEFQYAQSKPKVEELNVICNVSQFSVNK
PKCYKGVAFPLFTDAGVATATOTKSATINYHEWQVGASLSYRLNSLVPYIGVOWSRATFD
ADNIRIAQPKLFTAVLNLTAWNPSLLGNATALSTTDSFSDFMQIVSCQINKFKSRKACGV
TVGATLVDADKWSLTAEARLINERAAHVSGQFRF

781703 782599 KVB-010-00-00 ARTHOUGH AND ARTHUR HAND THE AND A LA ENT LE EMBETTATE LAGLAR REPARADORDA E GODDINE DE ENBERGOTT VERBENDE EN

782587 783447

ETYREE LAAKTDAMSTALLEVNVETDEVANNAVEREEVENLLIND HASVALAGACKKEHR WERKLMODESMETERTLEGGTEVELTEGKEALEAGASHLEEAVVYLRIKEGLAGAGKKEHR ETRE-ELORGATON FASTOL TEC SEQUESTION DELINATING TVOETH RESERVAY FEKATISTY OF YOR ON KEYALITMESCOS

TADSLAKDIAMHVVAAQPQFLSKESVPAEAIAKEKEV FFQEACLLEOPFIKNADLSIQSLIDDFSKTSGSSVAI

CKPOEV LEK IVTGKLNT

783443 784201 t*Pn 0648

CYP_DOWN

DYTH-UMP KINASE

EPNKNMAKOTRRVIFKIGGEALGKDSSNRIDEMRLSRLVSELRAVRNNDIEIALVIGGGN
LLRGLAEOKELDINRVSADOMGMLATLINGMAVADALKAEDIPCLLTSTLSCPQLADLYT
POKSIEALDOGKILCTTTTAGSPYLTTDTGAALPACELNVDVLIKATHHVDGVVDKDPRL
FFDALEYDEVGVERFELGEDE FYMDADALALERMEDHIEIRVFEFDJADDEKFALFDPTITTL SORDAMHACOURRE

784179 784721 CPn 0699

Trf-Ribosome Releasing Factor
TMSVLQDTEKKMAAALDFFHKEVKSFRTGKAHPALVETVVVDVYGTTMRLSDIASISVAD
LRQLVISPYDGNNASAIAKGIIAANLNLQPEVEGSIIRIKVPEPTADVRQEMIKQLRRKC
EEAKINVRNIRREANDKLKKDSALTEDVVKGNEKKIQELTDKFCKQLDELTKQKEAEIAS

CPn_0700 785094 785609 CT676 hypothetical protein LMYHSPTHCCYHCQOPATICYTEIDKDKVIRSYVCATCPCPSHYYNNEHLSLSKGVGVLT LECGNCKTVWHSKQDDEQLLGCHCCYTNFKNQITSKLKSERVVSSSFTMEKGQGSLHIGR APGEASMTNPLLKLIALNEALQDTLEREDYEQAAVIRDQINHLKTKNPDDPS

785584 786672 CPn 0701

CPH_0701 785584 786672
karg-arginine Kinase
KPKIQMTLPMDLLETLVKRKESPQANKVWPVTTFSLARNLSVSKFLPCLSKEQKLEILQF
ITSHFNHIEGFGEFIVLPLKDTPLWQKEFLLEHFLLPYDLVONPBGEALVVSRSGDFLAA
INFQDHLVHGIDFQGNVEKTLDQLVQLDSYLHSKLSFAFSSEFGFLTTNPKNCGTGLKS
CFLHIPALLYSKEFTNLIDEEVEIITSSLLLGVTGFFGNIVVLSNRCSLGLTEELLLSS
LRITASKLSVAEVAAKKRLSEENSGDLKNLILRSLGLLTHSCOLELKETLDALSWIQLGI
DLGLIKVTENHPLWNPLFWQIRRAHLALQKQAEDSRDLQKDTISHLRASVLKELTKGLSP
ESF

CPn_0702 789700 786929

yscC/gspD-yop C/Gen Secretion Protein D

LKKNPVKTVILNIGRKILQGIKKKKKKIGILSGLFFLDVLLGVSSQRPTETSANVKHNL
RDEKLAACPRNSAASLSAKKSHTKKTTPGSIPSKTPSKTPATQDKTFQKTSGSAFPAKPT

TLKELEERKKPREPRRTTADVKRSPFLPTOEVEEPVPAASKROLDSIQWEEKQNYARR

AVNAINLSIKKQLEEQTSTVTEKDVQPKTQATPHASKKNVASPSTSMEGIEKAATTVAVP

QDKSEEKVKERLTKRELTCEDLKDNGYTVHFEDISILELLQFVSKISGTNFVFDSDLQ

FÖWTIVSHDPTSVDDLSTILLQVLKHHDLKVVEGCNNVLIYBNPHLSKLSTVVTDSSLKE

TČĒAVVVTRVFRLYSVSPSAAVNIIQPLLSHDAIVSASSATRHVIISDIACNVDKVSDLL

AĀLDCRTSVDMTEYEVKYANPAALVSVCQDVLGTLAEDDAFQMFIQPGTNKIFVVSSPR

LAMKAEQLLKSLDVPEMATHLDDPASTALALGGTGTTSPKSLRFFMYKLKYQNGEVIANA

LQĐIGYNLYVTTAMDEDFINTLNSIQWLEVNNSIVIIGNGGNVDRVIGLLMGLDLPPKOV

YIEVLILDTSLEKSWPFGQWAALGDEGSKVAYASGLLNNGIATPTKATVPPGTPNPGS YIEVLILDTSLEKSWDFGVOWVALGDEQSKVAYASGLLNNYGIATPTKATVPFGTPNGS
IPHETPOQLTGFSDMLNSSSAFGLGIIGKULSHKGKSFLTLGGLLSALDQDGDTVIVLNF
RIMAQDTQQASFPVGQTVPYGTTNTIIQFEGTVTONIDVEDIGVNLVVTSTVAPNNVTLI
QIEOTISELHSASGSLTPVTDKTYAATRLQIPDGCFLVMSGHIRDKTTKVVSGVPLLNSI
PLEMBGLFSRTIDGRQKRNIMMTIKPKVISSFEEGTRVTNKEGYRYNWEADEGSMQVAPRH
APECCGPPSLQAESDFKIIEIEAQ

CPG_0703 791205 789685
pki6-5/T Protein Kinase
RKIGFMCRGGIPLPERQVICGYHVKKILSKKLRSRVVHGLHPETRHSTVIKVFSP9ESF
TÄRSVYNFLKEAGSLHOLTHPNIVKPHRYGKWODCLYIAMEYIEGISLREYILAQFKSLP
OAIDIIFDIAQALEHLHSRNILHKDIKPENILITPQGKIKLIDFGLADWDTEIQRAHPSV
IGTPYYMSPEQRGGESHSPASDIYALGLLAYELILGHLSLGRVFLSVUPERISKLAKAL
QSSPNNRYSTREFIQDIHHYRMSGDMQDLRIKDHTVALVEQLQTQRFWLAPFLIRFPD
FISGVIYHQGYPLYPHAYDTLLEGDVFNLWLGYSPISNATIALSVVKSLVCQDLQRPLL
DRYGEINECLIRMKIPIDEMGISILCLEISKENKELSMIAGGKTVFWIKRQGVVQDFES
FSPGLGKITSLQIRETKVAWEIGDEAVVCTLELEESVASLKTLSLAELQDRQCKAIFCPI
ESEHGGIQSRQHGSNSPSTLISLKRIR

CPA 0704 792330 791209

flin-flagellar Motor Switch Domain/YscQ family

RYÉMAVANDSSASWLKSRNNFLSSLGKTEEQVARPEFRELCOHKIJEKFRLEDVOVSIK

FRGSITAVEATKEEGVHLLIOPMVOPMEVENLLFLTSEEDLOELWAVFDDASLASYFY

EKDKLLGFHYYFVAEACKLFEELOWYPSLSAKVCGDAIFTATSLOSSGVVDISLRLDGK

NVRCRLLLPEDTFOSCQKFFSGLHDESDLHNIDOTOQISLSVEVSYSQLTOEEWHOVVPG

SFIMLDSCLYDPETEESGALLTVQKHQFFGGRFLTPSSGEFKIJSYPNLTHEDPPLPENP

QASAAPLPGYSRLVVEVARYSLAVSEFIKLNLGSILSLGNHPAYGVDIILDGAKVGRGEI IALGDVLG IRVLEV

CPn_0705 793176 792334
CT671 hypothetical protein
FMELKKTAESLYSAKTONHTVYQNSPEPRDSRDVKVFSYEGKQTRQEKTTSSKGNTRTES RKFADEEKRVDDEIAEVGSKEEEQESQEFCLAENAFAOMSLIDIAAAGSAEAVVEVAPIA VSSIDTOWIENIILSTVESMVISEINGEQLVELVLDASSSVPEAFVGANLTLVQSGQDLS vkfssfydatomaeaadlytnnpsolsslysalkgholtlkefsvönllvolpkieevot plhmiastirhreekdordonokokoddkeodsyk eearl

CPn_0706 793689 793180
CT670 hypothetical protein
YAVAKYPLEPVLAIKKDRVDRAEKVVKEKRRILETEQEKLREKEAERDKVKNHYMQKIQO
LROLLOEGTTSDAVLQIKSYTKVVAVQLSEEÉEKVNKQKEVVLAASKELEKAEVNLAKRR
KEEEKTRLHKEEWMKEALKEEARAEEKEQDEMGQLLFQLRQKKKRESGGS

7\$3704

CPn_0707 795035 793704
ypon 70p N (Flagellar Type y Phase)
ymmolittopditimgoladunutityvortitevyokutkavypnyrygevolukrnomepl
ytevyoptoptapisoladunutityvortitevyokutkavypnyrygevolukrnomepl
ytevyoptoptapisoladunutityvortitevyokutkokutkavypnyrygepidyet
ygerijityyptopterapodelirakorolustivikolokutkokutologepidyet
ygerijityyptoptapisoladunutaliteraevyepi egodijedejmeri vyvotodojsolokun
aayyotataeyprodoktyvimmosytraraliryglaavepparayytedypstipri
leeryaadokstytiapytolokutopyneeyyadeykstilothyvimalaavayhypatovla
disklutatyperakritikalirygepaavolikyterakykanemuterigepprodoketofatohidkurk
elkolothektiyperaavolikyter FLKOD HERTNYREAAOOLRAYLFR

CPn_0708

7/57 18 795034 CT668 hypothetical AFKTVKRFFCFMIDPVECF DAEAQ / ITONOGTELAGELKKO COPFALGSYAAPKD AFKTVKKEFCFHIDDVBEFF WOODANGT LONG FEAGUSTRAD GOPFAGGSYAAPKD TTLVOGFKENPAMODONGHILIDPELOBALEGEELDEQ INNLKGRLWEFFSTFEDSOTT AOFADEHFQAVGY I TOL INEDLINT I AEHTOODARKEDKEEGGSVTRK I IDWYSSGEEVLIN ALLYFSDRDGNRESLANFLK/QYAVQRATQRAELFAG I VOTSVSSVKT IMTTQLG

795203 CPn_0709 796203 CT667 hypothetical prote

EMPROPERTABILITY OF A STANFORM THE CHARLES AND A STANFORM OF THE STANFORM PROTECTION OF THE STANFORM OF THE ST VERRIMADARVE, ALFA

QVIANCKIESTRALAQSVLLWHOTLVAKSAUPLU

798482 796210

CT666 hypothetical protein RSRGERSMATNRSCTAFDSNKHLDGVCTYVKGVQQYLTELETSTQGTVDLGTMFNLQFRM QILSQYMESVSNILTAVNYTEMITMARAVKGS

796791

CPn_0711 / 795791 796486 CT665 hypothetigal protein TIINNQVLGFINYLYLGRYSHFINMENTAKEEKNSQPLLDLEQDMQDHDRAQELKASVQDK VHKLHALLREGSDIJESFGQQQSLLAGYVALQKVLGRINRKMI

799315 796781

THA domain: homology to adenylate cyclase)
MAURLIVDEGPLSG/IFVLEDGISWSIGRDSSANDIPIEDPKLGASQAIINKTDGSYYIT MAVRLIVDEGFLSG/IFVLEDGISWSIGRDSSANDIPIEDPKLGASQAIINKTDGSYYIT
NLDDTIPIVVNSVAIQETTQLKNEDTIILLGSNQYSFLEDFFDQDLYDFDIPEENFSND
SGDLSDSNGGKDLEPROTSETNINSPKPKEKLKTKDQGSSDPTTSGQOELADAFLASAKAE
KNQPRAKVAKKGLKESSNESLNPKEQNAKDSPKGEERTNKPQNAIMEDNGASPRQDPQPK
SAEPSLKYTARDETPLKENKPVEEKANKKATPDSPEKKOPEEGSKKEGSKIEATPLDSQ
KESEDKAEEAFVQEEEENLTEDNKEDSDSAADANDDTASDHATAEDNKETPKKVENEKSA
VLSPPLVQDLFRFDQTIFPAEIDDIAKKNISVDLTQPSFFLLKVLAGANIGAEFHLDSGK
TYILDTDFTTCDIVFNDLSVSHQHAKITVGRCGILIEDLDSKNGVIVEGRKIDKTSTLS
SNQVALGTTIFFLIDHHAPADTIVASLSPDDYSLFGRQQDAEALERQEAQEEEEKQKRA
TLPAGSFILTLFVGGLAILFGIGTASLFHTKEVVPLENIDVQEDLAQVINQFFTVRYTFN
KTRSQLFLIGHVKNSTDKSELLYKVDALSFVKSVDDNVIDDEAVNQENNILLSKRPEFKK
TSMSGLFLIGHVKNSTDKSELLYKVDALSFVKSVDDNVIDDEAVNQENNILLSKRPEFKG
ISMHSPERGKFIITGYVKTEEQAACLVDYLNIHFNYLSLLENKVVVETQMLKAIAGHLLQ
GFANIHVAFVNGEVILTGYVNNDDAEKFRAVVQELSGIPGVRLVKNFAVLLPAEEGIID
LNLRYPNRYRVTGYSRYGEISINVVNNGRILTRGDVIDGMTVTSIQPNAIFLEKEGLKYK
LDYNK

799332 799817 CPn 0713

CMT_U/13 / 7931/1
CT663 hypothetical protein
LDLKERKAGFRNEIVSIPGOTKTIXALENTSMLEKLIKNFATYMGITSTLELDADGAYV
LPISEVVKVRAQQNADNEIVLSASLGALPPSADTAKLYLQMMIGNLFGRETGGSALGLDS
EGMYVMYRRSGDTTYDDFVRHVESFMNFSETWLSDLGLGKQ

801125 800091

CPn_0714 801125 800091
hemA-Glutamyl tRNA Reductase
nyRIVLMYLGYVGISYREAALKERERAIQYLQSFEKNLFLAQRFLGKGGAFIPLLTCHRA
ELYYYSESPEIAQAALLSELTSQGIRPYRHRGLSCFTHLFQVTSGIDSLIFGETEIQGGV
KRAYLKGSKERLEFDLHFLFQKALKEGKEYRSRIGFPDHQVTIESVVQEILLSYDKSIV
TNFLFVGYSDINRKVAAYLYQHGYHRITFCSRQQVTAPYRTLSRETLSFRQPYDVIFFGS
SESASQFSDLSCESLASIPKRIVFDFNVPRTFLWKETPTGFVYLDIDFISECVQKRLQCT
KEGVNKAKLLLTCAAKKQWEIYEKKSSHITQRQISSPRIPSVLSY

CPn_0715 801636 803462

gyrb-DNA Gyrase Subunit B
KFNKISHMAAYTEASILSLASLDHIRLRAGMYIGRLGNGSQKEDGIYTLFKEVVDNGIDE
FINCHGKSLKISASDKQISIQDQGRGIPLGKLIDCVSKINTGAKYTODVFHFSVGLMCVC
LKAVNALSEIFSVRSVRKKYHLATFHRGVLQESKQGSTKDPCGTFVSFTPDPSIFPEFT
FNHDFLKDKIRQYTYLHSGLEIRFNDEVFISHNGLKDLFDAEITEPPLYSPLFFQNEDLT
FIFSHLEGNTERYFSFNAQGETLOGGTHLTAFKEAIVKGVNEFFGKTFYSNDIREGIVGC
IAIKIASPIFESQTKNKLGNTQIRSSLIKDVKEAIVAGVNEFFGKTFYSNDIREGIVGC
FILTGAVESLGKEKPMNYFSLEETKMYKNDELFYLATALGITONEIGHLRYMVILATDAD
DGMHIRNLLITFFLKTLLPLVENNHLFILETPLFKVRNKTTTLYYYSEQEKMQALQOFGK
KDSSLEITRFKGLGEISPKEFAAFIGPEIRLTPVTITSLESISSILQFYMGKNTKERKQF
IMDNLITDF IMDNLITOF

CPn_0716 803466 804902

gyra-DNA Gyrase Subunit A

FMRDVSELFRTHFMHYASYVILERAIPHILDGLKPVQRRLLWTLFLMDDGKMHKVANIAG

RTMALHPHGDAPIVEALVVLANKGYLIDTQGNFGMPLTGDPHAAARYIEARLSPLARETL

FNTDLIAFHDSYDGREKEPDILPAKLPVLLLHGVDGIAVGMTTKIF PHNFAELLKAQIAI

LNOKKFTVPPDFPSGALMDPSEYQDGLGSITHLASIDIINDKTLVVKGICPGSTTETLIR

SIENAAKRGTIKIDTIQDFSTDVPHIEIKLPKGSRAKEMLPLLFEHTEGQVILYSKPTVI YENKPVECSISEILKLHTTALQOYLEKELLLLQEQLTLDHYHKTLEYIFIKHKLYDSVRE VLAINKKISADDLHQAVLHALEPWIHELATPVTKODTSQLASLTIKKILCFNEEACTKEL LAIEKKQAAIQKDLGRIKEVTVKYLKGLLERHGHLGEPKTQITNFKTAKTSILKQQTLI

804968 805306

CT556 hypothetical protein
IRIKFIDTITIVVRMEPRHIYIRKPETPKAPDVEKPGVPEYMTMANTPTFEGPVKTLDQL
RRALIEQRGAEEGYMYDNFIQSILISTFGLVHKDMDPAQKASKRMRSVYKEQ

805300 805626

CT557 hypothetical protein RAVMSFTYFLALPYDRIMGERFLCSPKRWAPFINSPLYLTLIADHDTRYLAKNLDKFPLP VECWEKTYLHYSSLIKSIFLSSDLSSLRLLACTKFEILTLHDLYCAQNI

R05877 806890

CPA_9714 R05577 B06840

STAB (Pseudour Lifton Synthème)

FDIPLYKKYTK REMETVESFTYCKENCORLDKYLTEVHPKYCRAFYQERTILOGLVQINO

HIPPAVARIACONTTIDIÇEREELLELLIPATI-LDPY/ECOMITATIKRENMAVHAPG

HEPFOTANIALMELGERIKEEFPEREMRISTIVHRLDFY/ECOMITATIKRAKKYFSELFS

TPFELKESYLAVETGKIROTETHTHI GRHQNKKKEMTYZOZ/KEAVPHCQVIAFROKLOFY

ALDESTGRENGLEVIMKHLATETIVIDDVYGTPSMNCOYGLDNOOLIAYSVUFTHPETROF CORPORTEGE ORDER WHEN THE TRANSLESS TEKES

CORPORTEGE ORDER WHEN THE TRANSLESS TEKES

CORPORTEGE ORDER WHEN THE TRANSLESS TEKES

CORPORTEGE ORDER WHEN THE TRANSLESS TEKES TEKES

CORPORTEGE ORDER WHEN THE TRANSLESS TEKES

PDICKLICKEGRTIKAIR

807571 808489

KdSA-KDO SYNTHELISE KRMMFNNKMILIAGECYLEGEDITLEIAGKLQSILAPYSDRIQWFFKSSYDKANRSSLN GFROPGLTEGLRILAKVKETFGVGILTDVHTPQDAYAAAEVCNILQYPAFLCRQTDLLVA TAETMAIVNLKKOOFLSPWDMEGPIHIYJLSTCHRIKILLTERGCSFGYNHJVSDMRSIPVL GROGEFYLEFATHLYYLLKIALSTCHRIKTERGYTCLGRAADAW DAKCH LETHETHEFTAKS DAASMLJLEEPAALLUWWLJGFFTZCGEMICA

808477 308974 CPn 0722

CYN_UV22
CT654 hypothetical protein
YGLSMTKFLYCGLFYSLGLLVLAFGTMVAIIQVDQICDVSCMNKHFQESPPFLKIKKVNV
SKQICSPEERFHCKINGSCMELHFPQSSYSCKEYLTRISGHILTQNFEKQMQFRGNSGL
LNYQDGSLHVYDCRFQVDPVPGYGSPDKEDSSSGGMKTLYLSLFRN

CPn_0723 808978 809703
yhbg-ABC Transporter ATPase
ASMPILSVCNLVKKYNKKPVTNDVSFQINPGEIVGLLGPNGAGKTTAFYLTVGLIRPDSG
KIIFKAVDVTKKTMDKRALGIGYLAQEPTIFKELTVQDNLICILEIIYKARKQQSHLLN
TLVDDLQLGSCHKKAGTLSGGERRRLEIACVLALNPSVLLLDEPFANVDPLVIQNVKYL
IKILAGRGIGILITDHNAKELLSIADRCYLIIDGKIFFEGSSSQMISNPMVKQHYLGDSF 808978 809703

809706 810602 CPT_0724 810602 899706
No robust homolog present in Genebank/EMBL as of 11/7/98
RTSTRLDYRSGCILSKILPFPELWKMLLGFLCDCPCASWQCAAVANCYDSVFMSRPEHKP
NIPYITKATTRGLRMKTLAYLASLKDARQLAYDFLKDRGSLARLAKALIAFKEALQEGNL
FFYGCSNIEDILEEMRRPHRILLLGFSYCQKPKACPEGRFFDDACRYDFSHPFCASCSIGT
MMRLNARRYTTVIIPFFIDIAKHLHTLKKRYPGYQILFAVTACELSLKMFGDYASVMNLK
GVGIRLTGRICNTFKAFKLAERGVKPGVTILEEDGFEVLARILTEYSSAPFPRDFCEIH

CPn_0725 810829 810587
CT652.1 hypothetical protein
SCGDVGMFFAPLLYESLRRGLMHPTSHMQQQLARLEFINDQLTTELEHVNELLCSLGFPE
GLTTIKAIAEEVLSDDEPLLD

813384 810880

CPI_0727 813559 816192

CPi_0727 816192

CP

CPn_0728 818483 816525
CHLPN 76kDa Homolog (CT622)
VPM/NPIGPOPIDETERTPPADLSAQGLEASAANKSAEAQRIAGAEAKPKESKTDSVERW
SILRSAVNALMSLADKLGIASSNSSSSTSRSADVDSTTATA/PPPPFTFDDYKTOAOTAY
DTIFTSTSLADIOAALVSLODAVTNIKDTAATDEETAIAA/WETKNADAVKVGAQITELA
KYASDNOA ILDSLGKLTSFDLLQAALLQSVANNKAAELIKEMQONPVVPGKTPAIAQSI
VDQTDATATOIEKGGNAIRDAYFAGONASGAVENAKSNSISNIDSAKAAIATAKTQIAE
AQKKFPDSPILQEAEQMVIQAEKDLKNIKPADGSDVPNGTTVCGSKQQGSSIGSIRVSM
LLDDAENETASILMSGFROMIHMFNTENPDSQAAQGELAAQARAAKAAGDBAAAALADA
QKALEAALGKACQQQGILNALQGIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDY
KTQISAGVDAYKSINDAYGRARNDATROVINNVST/ALTRSVPRARTEARQPERTDOALAR
RVIGGNSRTLGDVYSQVSALQSVMQIIQSNPQANWEEIRQKLTSAVTKPPQFGYPYVQLS
NDSTQKFIAKLESLFAEGSRTAAEIKALSFETNYLFIQQVLVNIGSLYSGYLQ 818483 816525

CPn_0729 819905 CHLPN 76kDa Homolog (CT623) R1859 CHLPN 70KD HOMOTOT (CT023)
PAWSCVETUNIOTKOTMKKQVYQWLASVVLLALTISGYAELFLGEQNVKSHTYTTLDEVK
DYLSKRGFVETKKODOVLR LACDVRARMLYFRED IKHPSDKDKYNPLPVNRYRSEFYLYI
DYRAERIMLSSKMMMTA LAKSENTAAGVYINRAFLGYRFYKNPETRTDFFME IGRSGLGD DYNARIAMIJSKAMWYA LAGGENTAA WY LINAFEG YR Y KREETKTEFFE LAGOKAP OFFVKODVUMTEVPOET: TTEKANTMAMKYKY UWWILVIKHENYAWUSI LAGOKKPLYLY OFFVKODVUMTEVPOET: TTEKANTMAMKYKY UWWILVIKHENYAWUSI LAGOKKPLYLY ORGALLIKHENGAQI AANNOPKEAN FTDY LIKKO I WOTTUGISFRAYGAYOKPANDK LOODPTEKKEDIATI COF

CPH_0740 821676 819963 mvin Indeshat Membrany Protein CCPKGRGENGEMERKEMEVSTARSTENILEXTPCCFPTGTFRETAMATYFGADDIVAAFW

LISERTVEFLERILGSLILE

AFEFLRAGSCRAAFFERREGRLIKGST. IFTLLIE

AVLAVVLOYVEEGTYDMILLS. LLPCGIFLMYNMOGALLHCENKFF7YGLAPVVNII

WIFFVIAARHSDPRER! IGLSVALVIGEFFEYLITVEGVWKFLLEAKGPFQEHDSVRALL

APLSIGILTSSIFOLILLSDIGLARYVHEIGFLYLMYSIKIYQLPIHLFGFGVFTVLLPA

ISRCVQREDHERGIKLMKFVLTLTMSVMI IMTAGLLLLALPGVRVLYEHGLFCSAVYL

URVLRGYGASI IPMALAPUVSVLFYAQRQYAVPLEIGIGTALANIVLSLVLGRWYLKDVS

GIGYATSITAMVQLYFLWYYSSKRLPMYKKLWESIRRSIKVMGTTMLACMITLGLNILT

OTTYVIELNPLTPLAMPLGSITAQAIAYLGESCIFLAFLFGFAKLLRVEDLINLASFEYW

BERGERLAGARERM 144.6 H7M-

CPn_0731 821474 821760
No robust homolog present in Genebank/EMBL as of 11/7/98
VALAISRNIPVIRLOM/PDNILK/ERAKETSLSFLLIKPFSPPPLKQDYLFDISPYTSSE
ITIGGSYFKLNKASLQSSTLRLKSISIIS

822/092 CPn 0732

CPH_0712 822092 622770
nfo-Endonuclease IV
nfmkvlpppsipllgahtsragglknaiyegrdigastvqiftanqrowqrralkeevie
pfkaalketdlsyinshagylinpdapppvilleksrigivqeildcitlgisfvnfhpga
alksskedcmnkivssfosaplfpsspplvvllettagottligsnfeelgylvonlkn
Qipigvcvbtchifaagyditspqgwedvlnefdeyvglsylrafhlndsmfplgankdr
haplgegyigkesfkylmtdertrkipkyletpogpenwqkeigellkfsknrds 823739 823101

CPn_0733 CPM_U733 823739 823101
rs4-54 Ribosomal Protein
GLKYMARYCGPKYMVARRFGANIFGRSRNPLLKKPHPPGQHGMQRKKKSDYGLQLEEKQK
LKACYGMIMEKQLVKAFKEVIHKQGNVAQMFLERFECRLDNMVYRMGFAKTIFAAQQLVA
HGHILVNGRRYDRRSFFLRPGMQISLKEKSKRLQSVKDALESKDESSLPSYISLDKTGFK
GELLVSPEQDQIEAQLPLPINISVVCEFLSHRT

823863 824915 CPn_0734

YCEA
ONTKEHFSSINGNFLOCHYFODYVRVFIMEKKYYALAYYYITRVDNPHEEIALHKKFLEDL
DVSCRYISEGGINGOFSGYEPHAELYMOWLKERENFSKIKFKIHHIKENIFFRITVKYR
KELANLGCEVDLSKOAKHISPOEMHEKLOENRCLILDVRINN EMKIGHFINATLPDIQTF
REFBEYAEKLAQECDPETTPMMYCTGGIRCELYSPVLLEKGFKEVYQLDGGYIAYGOG
GTGKMLGKLFYFDDRIAIPIDESDPDVAPIAECCHCQTFSDAYYNCANTDCNALFLCCDE
CIMONGCCGEECSQSPRVRKFDSSRGNKPFRRAHLCEISENSESASCCLI

Pn_0735 825680 825003

*Uridine Kinase (Uridine Monophosphokinase) (Pyrimidine Ribonucleoside Kinase).

GEKFMLMMLMMIGITGGSGACKTTLTONIKEIFGEDVSVICQDNYYKDRSHYTPEERAN LIWDHPDAFDDLLISDIKRIKNNEIVQAPVFDFVLGNRSKTEIETIYPSKVILVEGILV FENQELRDLMDIRIFVDTDADERILRRMVRDVQEQGDSVDCIMSRYLSMVKPMHEKFIEP TRKYADIIVHGNYRQNVVTNILSQKIKNHLENALESDETYYMVNSK

CPn_0736 827731 825992
ygeD-Efflux Protein
RGELLKLARQCLVAFMTVSVKKKSFRALVTTHFLTIINDNLYKFLLAFFLLEGKTLTENA
KILSCVSFFFALPFLLLAPLAGSLADRFQKRNIILATRFIEILCTILGTYFFFIQSVVGG
YVVLILMACHTTIFGPAKLGILPENLPSEQLSQANGIMTAATYTGSILGSCLAPLLVDVT
HRLGVNSVWPTLMCVTVSIISTLISFCIRPSNVKNVKQKITLVSFKDLMKVLKDTRMIH
YLTVSIFLGSFFLLIGAYTQLEIIPFVEFTLKYPKHYGGYLPFIVALGVGTGSYITGKIS
GKDIKIGYVPLAAIGLALVFMGLYAFACSILFVLFFLLALGFLGGVYQVPLHAYVQYASP
EHKRGQILAANNFLDFFOULVAAGVIRVLGSNLGLSPETSFFYIGMFVLAVSIWTLWIMR
EHVYRLLIGIILRRQLGYYLKIHOSSSPKCYFVAVQSYREIRRVLAALTKTVRSRVIILD
QKLVPGWRAMLLSWCVPTVVSSVRDNDSEAQDAWAVLQANHLKTSLKKFPDVSVVCLGLP
KNVERFFSILQEGGIDLHPIOLVQKECKKRVIYTLVFPHA 827731 825992

KNVERFTSILQEQGIDLHPIQLVQKEGKKRVIYTLVFPHA

CPn_0737 827469 830756

"recC-Exodeoxyribonuclease v, Gamma"

KRSAKLPASGASKRKGRAKKKLTQERIFAFSVRVLPSNRKNAKRNLYKLSFIIVRKCVVT
SALNDFFLTETVMNATKHCRASFSNSPRHLLAQLAEDITSTHQKPFTKRWILVANATTGH
WIKNQLVMVLSDHIFMGSTIFTASDSIVKHLFLGSGCSQPNT PDYVLTLPLLINNILEEIS
KASKFENGREFLSPFTYETTKKLAAAFKOFHTFSQRPTKNASHYQELFQILESHFSSYEE
MFTTILNNRTQEEDCSLHIFGYAHLPKHLAEFF INLSTYFPVYFYCFSFCEYFGDLLS
KASKFENGREFLSPFTYETTKKLAAAFKOFHTFSQRPTKNASHYQELFQILESHFSSYEE
MFTTILNNRTQEEDCSLHIFGYAHLPKHLAEFF INLSTYFPVYFYCFSFCEYFGDLLS
KASKFENGREFLSPFTYSTTKKLAAAFKOFHTFSQRPTKNASHYQELFQIVGEMFLPSK
HDSSLGVIQNSILDLKPTSPODFSGTKGTICIYRALNIFREVQEVFCKVTELHARGVSPE
EIFILSSHIESYKVHLNAIFFNHPVPIYFTDEVDPRAEDLRNKILLLSSILQTQGDLHYIL
OLLTHPQLQQPIDONKVPYLIKKLSSENGKISSKDRASGQOMKALGDLILEEYPFHQEGG
RVSQVEVWETTVPLIYFIQERINLYLSSSQHSYEDLFQNVTSCLEKIFYLSPETSFITL
LGRNKTTSSDIFDLLNRTTTHEELAFSSTEDEENFHFLOILVSTKHELHISYISSAQFN
LGRNKTTSSDIFDLLNRTTTHEELAFSSTEDEENFHFLOILVSTKHELHISYISSAQFN
LPSPFLNHIKETLOLPVETLPFTQPYLSAFFKNKACLHTSGEVTYNSLAHAFYSKKALLPSL
FIPTVKQVNLPPHLSLNEIIKGIFSPLDLFLKTTNNLRISYPEHLKKQCKJFTKHQIED
FWNECFVDKEHDLIFSISPHAEELFTYYREKTILLRNGLDKDPKHSPYTYTFSSSIFED
ENYHECYLEPPLSLSFCONPVQIHGTIHCVCNEGLYLCSIDPRDSLKKTTRTLGSLPETSS
EQKPLENLSPLCWMTUDDEEKFHQAVLSAISEEAKNPSLPIFWQFHNRNIEEILNNVGAS
ERLKILSLFRGPCEAV ERLKILSLFRGPCEAV

CPn_0738 830719 833895

"zecB-Ekodoxyribonuclease V, Beta"

KPYLFSEVPYKPPNIFDSNSSIGKFFLEASAGTCKTFTIEQIVLRALIEGSLTHVEHAL AITEFNASTNELKVRIKDNLSGGKFFLEASAGTCKTFTIEQIVLRALIEGSLTHVEHAL AITEFNASTNELKVRIKDNLAOTLRELKAVLNSOPASLPTYLDINCNVKQIYMQVRNALA TLDOMSLETINGECNFYLEOYFPKTBLIHKNPALTHSQLVLHITNYLKQDLWKNVLFGE GPHLLAVRYNTTSHKTSSLVKKLLAGYTOPICSYFGSKYERLEOIGLAMMQV INSSLLET KOVFLDQUTAHISGFKKQPFSILDDLHHFVDLLYTSETHSSLFSFFKIAETNFKHRLAR YKRCAAFTVLENMENDETEFCHLDER FHTLLYDDLGYTLKONYTTPMLSPDESVFALEKL LSSCSAGPOVQALERQYQLVLIDFFODTDKOVMSIFSHLETISFKFTGSSFLIODFKQSIY EMRJADLFTYLTAKSSESEDKOLGLVTNYRGTFHLMBALIDGIFGKISFLETBOPKQSIY EMRJADLFTYLTAKSSESEDKOLGLVTNYRGTFHLMBALIDGIFGKISFLETBOPKQSIY EMRJADLFTYLTAKSSESEDKOLGLVTNYRGTFHLMBALIDGIFGKISFLETBOSLFBLISFY HALHPOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFETBOSLFFFTBOSLFFETBOSLFFFTBOSLFFTBOSLFFFTBOSLFTBOSLFTBOSLFTBOSLFTBOSLFTBOSLFTBOSLFFTBOSLF 830719 833895

GV 1F 1RG 1DTQCINGFFALNSSED 1PNFNPKA 1QKCQA

CPn_0739 R34892 833861

CT 368 hypothetical protein CKVLFKLMSYSLRNKYTKICVYIIIALGILSFRSIPQEVYDKIRSSFVSLHVKFFPKIKQ APSSHLANLELENLVLKERVASLEEKLKLYEVSNHTPPLFPEILTPYFHKLVEGKVVYRD YTHWSSSCWVNVGKTHGIKKNSPVLSGNVLVGLVDYVGEHQSRIRLITDVGMKPSVVAMR OD LOSSWEITERGEREELTPOVEOTSHAYTEERDRYEKTSOLOEELDSELTOOEGENOALERGTE PROPERATIONEELD IN THE RESERVE DER IKTELIER HELD TESTED VERFERDLINARVTEVEARED SACTIVE FRANCHISCHER LIBERTARIE LEPTERFORTE DEFECTIVATION OF THE SACTIVE OF THE S

836054 834864 CPn_0740 CPG_0740 836054 834864

tyrB-Acomatic AA Aminotransferase

SYMSFFNHIPTFSPDAILGLQNVFFADKRPEKVNLVIGVYEHPQKRYGGLSCIRKAQTVI
LEEEQNKSYLPISGLOIFLDEWRELVFGAVDPSAIVGFOSLGGTGALHLGARLLSVAKGS

KVYVPEDVINSNHIRIFSQGGLEVIRYPYYSKEQNGLLFEPLIAFLKEVERNSVILLIAG

CHNPTGVDFTEDMWKELAILMERELIPFPDTAYQGFAHGIELDRKPIEIFISEGNTVLV

AASSSKNFALYGERVGYFAVHSTFTDELVKIHSFLEEKIRGEYSSPQRWGVEIVSTILSN

PYLKEENGSELNFIRESLGKMRTRFVQALRKVAGHTFDFLLSQHGFFAYPGFSDKQVLFL

REQHAVYTTAGGRMNLNGITEKNIDHVVQSFIQAYEL

CPn_0741 838383 836185
greA-Transcription Elongation Factor
EYIFRLKTODIVDYLEKLOVLIEEGGSANFLSLWEEYCFNDVVRGRELVEILEKVKSSSL
ASLFGKIVDTVVPLEKE IPEGRKDRVLQLILDLOTSNSOMFFDIATEYVNKKYSGEENF
NEALRVVGLRDGRDFQFSLSRFDFLMHMHKGNFVFHQGGWGVGEVMGVSFLQQKVLIEFE
GIMSAKDISFETAFKSLTPLSGDHFLSRRFGDPDGFEAFAKENFIEVVEILLRDLGPKTA
KEIKDELDULVIPEADWNRWMGSATKIKKGTRIISPDNPKEPFVVLSDAGCSHMGQLEKT
LGSLNSAEKISLIYHFIRDLHSELKNIEIRKSLVKALQDLDVEEGNKSLILQRELLLSE
VLGIKDASIDKEYITSLSEDDTSRLLENMPIVALQKSFLSLVKYSSFWQOVFMQILLYT
TSPTMRDFVYKTIKNDPSSVEVLKKRLLDSAHQPMMFPELFVWFFLKLGNHEDGLFDPED
KEVLRLFLESALNFMYQVASTPHKELGKKLHHYLVGQRYLAVRQMIEGASLPFLKELLLL
STKCPQFSSSDLNVLQSLAEVVOPTLKKHKSNVEEDNVLWSTSESFSRMKALQSLVGK
WYDNAKEIEDARSLGDLRENSEYKFALEKRARLQEEIRNARILTKUDLVFTDKV
GVGCKVTLKGDAGEVVEYTILGPWDADPDSCILSLQSKLAQNMLGKKLNDVVLLQGKEYK
ISRIQSIWEEHGA

CPn_0742 838442 838888
CT635 hypothetical procein
TKMMVIVMNSKSAQKIIDSIKQILTIYNIDFDPSFGSSLSSDSDADYEYLITKTOEKIQE
LDKRAQEILTOTGMSKEOMEYFANNPDNFSPEEWLALEKVRSSCDEYRKETENLINEITL DEHPTKESKRPKOKLSSTKKNKKKNWIPL

838956 CPA_0743 838956 840362

**ngkA-Ubiquinone Oxidoreductase, Alpha*

IPHKITVARGLDLSLOGSPKESGFYNKIDPEFVSIDLRPFQPLSLKLKVEQGDAVCSGAP
IAFKHFPNTYITSHVSGVVTAIRRGNRRSLLDVIIKKTPGPTSTEYTYDLQTLSRSDLS
EFFFENGLFALIKORPFDIPAIPTOTPRDVFINLADNRPFTPSPEKHLALFSSREEGFYV
FVAGVRAIAKLFGLRPHIVFRDRLTLPTQELKTIAHLHTVSGFFPSGSPSIHIHSVAPIT
NEKEVVFTLSFQDVLTIGHLFLKGRILHEQVTALAGTALKSSLRRVYITTKGASFSSLIN
LNDLSDNDTLISGOPLTGRLCKKEEEPFLGFRDHSISVLHNPTKRELFSFLRIGFNKPTF
TKTYLSGFFKKKRTYTNPDTNLHGETRPIIDTDIYDKVMPMRIPVVPLIKAVITKNFDLA
NELTFLEVCGEDFALPTLIDPSKTEMLTIVKESLIEYAKESGILTPHQD 840362

CPLD744 841387 840389
hemB-Porphobilinogen Synthase
emssltlsrrprrnrktaairdllaethlsprdliapffvkygnnikeeipslegvfrws
emssltlsrrprrnrktaairdllaethlsprdliapffvkygnnikeeipslegvfrws
emssltlsrrprrnrktaairdllaethlsprdliapfsvssprknilchsiheiknafphlcl
blyllferrctyflardefyllgevladesvrifgniatlhaemgadivapsdmdgrigyi
rskloosgysktsimsysvkvasclyspfroalsshvtsgdkkQygmnpknvlealless
ldeeegadilmykpaglyldviyrirgnviclplaayQvsgeyamilsafQQgwldketlf
HessitaikragadmiisysapfilellhQgfef

CPh 0745 841903 841742
No robust homolog present in Genebank/EMBL as of 11/7/98
VDSCFDDWRASSLOGSTTYNVAYDPKHTLAYGFCNQVSVKKFHLKPPKSQEKFL

CPAL 9746 841939 843567
CT632 hypothetical protein
FSGRCPFSFEVFMLGKEEFTCKOKOCLSHFVTNLTSDVFALKNLPEVVKGALGSKYSRS
VLGRALLLKEFLSNEEDODVCDEAYDFETDVOKAADFYORVLDNFGDDSVGELGGAHLA
MENVSILAAKVLEDAR HIGSSPLEKSTRYVYFDOKVRGEVLYYRDPILMTSAFKDMFLGTC
DFLFDTYSALIPQVRAYFEKLYPKDSKTPASAYATSLRAKVLDC IRGLLPATLITNLGFF
GNGRWONLIHKLOGHNLAELRRLGDESLTELMKVIPSFVSRAEPHHHHHQAMMOYRRAL
KEQLKGLAEQATFSEEMSSPSVOLVYGDPDGIYKVAAGFLEPFYSNRSLDDLIDVCKMH
HEDLVOILESSVSARENRRHKSPRGLECVEFGFDILADFGAYRDLORHRILTOERQLLST
HHGYNFFVELLDTPMEKSYREAMERANETYNEIVOEFFBEAQYMVPMAYNIRWFFHVNAR
ALCWICELRSQPOSHQNYRTIATGLVREVVKFNPMYELFFKFVDYSDIDLGRLNQEMRKE
PTT

CPn_0747 843949 844053 CT631 hypothetical protein RTCMGCKGAEVQILSSRSLSGMKILSSSLFYKKFC

844996 CPn_0748 844996 844121
ispA-Geranyl Transtransferase
GTLVLHALDTYRPSIESAIEKALEGFGPIGHPIRSPYEYALOGOGKRLRPGLVCMMAQGL
GLNNIDVMDSALAVEFYHTOTLIADDLPCMDNDDERFGREFVHKAFDEATALLAGYALIPA
AYSHLELNAKKLKEQGOGPREFIDIAVNIIGOTBOKNIGGGGVLCGGYDDMFFSNRGQEHV
QSIMIKKTGSLFETACISGWLFGGGDPGFAPIITSFSNAFGLLFQIKDDFSDLCKDSQQI
GLNYALLFGEKAALELLARSQNNCLELLDRLSAGGLKMSSEFETIISSLGSF

CPI_0749 84563H 845006
glidt-UDP-cHenAc Tyrophosphory Lase
VCYMPYLASE FSD-SDFLYFEL D:KAHYTWDTLDZMZ-MLENHVFSGTHGTVESGVTLKN
LEKTETAEDAYVESGAYTYCHS LLSSQ/TEVRIGAYLEGHVTTSTR:WVCHCTETKNSYLG
HETKAAHPAYUSESVALSEWNESGAVVEANFRUSPHTYVRSTSDKSKKIDTSBRKLGAF
LSKGVATGSNVVINGQHTLDHTRIBRSQVI

C1m_0750 846465 R45/07

onal Regulatory Protein - Receiver totD/cpxR-HTH Tran

Doman KITDFILRIHSYNLFCFHMIGDKIILFVTEDLSUSSQLKDLASORSDYQILVSPVFPTSF ESVAIFCEYLLLPEQIFSFSIFPEEDLIVLFDTFCEEAITKVLNCGATGYLLRPITAKVL DAVIRAFLROHEVLEHSIPDTMTFGDHTFRVLNLVIESPEGSVYLTPSEAGILKKLLINR GHLGLRKNLLAEIKGNTKEIIARNVDVHIASYRKKLGPYGSKIVTIRGVGYLFSDDDSIP GHLCLRKNLLAE II LONHONTAHPNEE

240 cost 848434

CTAGE MYSSING AGAIN STATE TO THE CONTROL OF THE CON YLYSMWLTKVAPSPOSYRLAGICLMENKRYDEALEFLCMLSPNDSINDYKTOKALAFCOK HQSKDRAAS

CPn_0752 448595 850082
recD-Exodeoxyribonuclease V. Alpha
GWALHTEFAPFLEDLVHOOVISPLDIAFASKHISSDFEESFVFLAVSSALWRYGHPFLSL
EENRIRPSLGGISETDLVRGFHNLENFEARDKLFVVVSGRLYLRSLYTIRSKLLDKLSLLC
SATPNYFPPSIDSSILSEEONFIFNKITOGCFSIVSGGPGTGKTFLAAQLILSLVKQOPK
LRIAIVSPTCKATSHIRQILHKYNIFDDMVLNGTVHHFLQEYAYRRYNSIDVLLVDEGSM
VTFDLLVSLVQTLGYYEMDKLYTSSLIILGDTNQLPPIGIGYGNELQDLIGYFHENTFF
LKTSHRAKTGVVDQLTQSVLRGEMISFSSPLPSISSAIEVLKNRFVKSLRQSEARLCVLTP
MRHCPMGVLLINMTHINQLAARSDPDLRIPIMVTSRYETMGLFRGDTGLLCLKTQKLHFPQ
HEPIDSRALSQVYNYVMSVHKSQGSEYDEVIVIIPKGSEVFGVSILYTAITRAKYRVSV
WGDPETLHKIJKKSNY

CPn_0753 851009 850161
NO robust homolog present in Genebank/EMBL as of 11/7/98
IMATAHLEROALLHIRSWTPAIRASGNLFRQOSMSLHNNVLFAGDIVGAIKNSTAISRHA
LGSSHYMAALQKTEGFLGAADGVNTAVAGAHLMGQLLNGSMIFETDEETGELRRCNEAD AEGCMYOKLORRSALTITGKVARLASKTLGTATFLHEMDVVSLGANANKIGCKVTSCLNL VATGCELTESSISLYRILSTRPETISDPENRNKPSAEFAARSKAIRNAFIAWLGDVVDLV CDALGTLSLFLPAILGVHAVLIMAILGLISCVINFVKDYAKIG

CPn_0754 851381 851040 rsk0-s20 ribosomal Protein OFILNLKVLVLSGDIMAPKKPNKKNVIQRRPSAEKRILTAQKRELINHSFKSKVKTIVKK FEASLKLDDTQATLSNLQSVYSVVDKAVKRGIFKDNKAARIKSKATLKVNARAS

851579 852799 (CPL_0755 851579 852799
TT616 hypothetical protein
YKDLFFMLLVRKWLHTCFKYWIYFLPVVTLLLPLVCYPFLSISQKIYGYFVFTTISSLGW
FFALRRRENOLKTAAVOLLOYKIRKLTENNEGLRQIRESLKEHQOESAQLQIOSQKLKNS
LFHLOGLLVKTKGEGQKLETLLLHATTEENRCLKGMOVDSLIGEGEKTEEVOTLNRELAET
LAYQOALNDEYOATFSEORNWLDKRQIYIGKLENKVQDLMYEIRNLLOLESDIAENIPSQ
ESNAVTGHISLOLSSELKKIAFKAENIEAASSLTASRYLHTDTSVHNYSLECROLFDSLR
ZENLGMLFVYARQSGRAVFAKALFKTWTGYCAEDFLKFGSDIVISGGKQMMEDLHSSREE
CSGRLVIKTKSRGHLPFRYCLMALNKGPLCYHVLGVLYPLHKEVLQS

CPI_0756 852889 854676

ITOD-RNA POlymerase Sigma-66
ISYLPLIKLSSKARNPLVLFOVRKLFMNTONSQATEVSSEEESQKKLEELVALAKEQGFI
TYEEINEILPMSFDTPEOIDQVLIFLTGMDIQVLNQIDVERQKEKKKEAKELEGLARRTE
GTPDDPVRMYLKEMGTVPLLTREEVEISKRIEKAQVQIERIILRERYSAKEAISIAHYL
ISGKERFDKIISEKEVEDKTHFLKLLPKLITLLKEEDTYLENLLLSLKQPDLSKQEAKKL
NDSLEKCRIRTQAYLRCFHCRHNVTEDFGEVVFKAYDSFLHLEQINDLKVRAERNKFAA
AKLAAAKRKLYKREVAAGRTLEEFKKDVRMLQRWMDKSQEAKKEMVESNLRLVISIAKKY
TNRGLSFLDLIQEGMMGLMKAVEKFEYRRGYKFSTYATWWIRQAVTRAIADQARTIRJV
HMIETINKVLRGAKKLMMETCKEPTPEELAEELGLTPDRVREIYKIAQHPISLQAEVGEG
SESSFGDFLEDTAVESPAEATGYSMLKDKMKEVLKTLTDRERRVLIHRFGLLDGKPKTLE
EVGSAFNVTRERIRQIEAKALRKMRHPIRSKQLRAFLDLLEEEKTGTSKVKSLKSK

CPn_0757 854709 855134
folx-Dihydroneopterin Aldolase
PCIKNIALVIAIERYOLIISKFRMMLFLGCSVEERHFKQPVLISVTFSYNEVPSACLSDK
LSDACCYLEVTSLIEEIANTKPYALIEHLANELFDSLVISFGDKASKIDLEVEKERPPVP

856434 CPD_0759 856434 855577
fola-Dihydrofolate Reductase
LLVRPVIPCNEENPLSVEMCKNPSVRGTYACDERGVTGLEGKLPWINTEDLOFFSETTOK
PPTVMGRKTWETLEPKYFVDRAVZVFCHEKROSVHGETWYPSLEEFLLLDLSSPTFLTGG
GELVGLFLENGTVPDFFTSHTKKEYACDTFFPLSLLETWTKTVLRDTQKTTTCYYENHHS 356997

CPn 0760 856.163 CTG11 bypotherical protein
RHOPKLCLEIPKRODEVTMKITTIKTERIYPTEDDLYGILEGGLPKLNEG:IVVITGKIVS
LGEGAVVELEKVGYDELIKQEADAYVETEKYGIYUTKKMGILIFGAGIDENVEGYEVLY
DCEGRALKMTYONLLDGESAAAYLGMGYDEQTITALIEEAFKITHEGGTAWNEFFTYNYGKP
DCEGRALKMTYONLLDGESAAAYLGMGYDEQTITALIEEAFKITHEGGTAWNEFFTYNYGKP
LAEDEDLYGDLLGGSAAAYLGMGYDEQTITALIEEAFKITHEGGTAWNEFTALFG
LAEDEDLYGDLLGGMAWETPALFG 857698 858375

CFN_07AL 837678 838375
CTh10 hypothetical protein
GIMTSWTELLDKQIEDQHMLKHEFYQRWSECKLEKQQLQAYAKDYYLHIKAFPCYLSALH
ARCDDLQIRRQILENLMDEEAGMPNHIDLWRQFALSLGVSEEELANHEFSQAAQDMVATF
RRLUDMPQLAVGLGALYYTEIQIPQVCVEKIRGLKEYFGVSARGYAYFTVHQEADIKHAS
EEKEMLQTLVGRENPDAVLQGSQEVLDTLWNFLSSFINSTEPCSCK

St. 16.79

TOWN ROAD COMMENTED TO DESCRIPTION OF THE EELLY FOR PERCENTION OF THE PROPERTY ATHEISTIKTGALSLIALGIKOVPKGRVIETFGPESSCKTTLATHIVANAQINGGVAAY
IDAEHALDPSYASLIGVNIDDLMISOPDGEBALSIAELLARSGAVDVIVIDSVAALVPK
SELEGDIGDVNVGLOARMASOALRKLTATLSRSOTCAVFINOIREKIGVSFGNPETTTGG
RALKPYSSIRLDIRRIGSIKGSDNSDIGNRIKVKVAKNKLAPPFRIAEFDTLFNBGISSA
GCILDLAVEYNILEKKGSVFNYQEKKLGOGREFVREELKRNRKLFEEIEKRIYDVIAANK TPSVHANETPQEVPAQTVEA

860520 859972 VgfA-Formyltetrahydrofolate Cycloligase
NFPMTDPKIEKSALRKLFISIRRDLSEERKHEASSAVASFVRSFSKESVVLSFVSFNHEI
DMCEANRILIOKCTLALPRIDGENYPVLIPSIDDLISVVHPKDPFSKOTPISSDKITHV
LVPGLAFDQOGYRLGYGHGFYDRWLAQHPYPSIRTIGIGYCEQKIDRLPQESHDIPLSQI

CPn 0764 861819 860524

CT648 hypothetical protein
GYKSMDIKKLFCLFLCSSLIAMSPIYGKTGDYEKLTLTGINIIDRNGLSETICSKEKLKK GYRSMDIKKLICHTELSSLIAMSPITGKTIGDYEKETETGINI IDENGISETICSKERIKK
YTKVDFLAPOPYGKVMRWYKNIKGONVSCLIAYHTNGGIKGVLECLINNRAYGRYREMHUN
GNIKIQAEVIGGIADLHPSAESGWLPDQTTFAYNDEGILEAAIVYEKGLLEGSSVYYHTN
GNIWKECPHKKGVPOKKFLTYTSSKLLKEGNYQGKRANGLSIRYSEDSEEDVLAWEEYH
EGRLLKAEYLDPGTHEIYATIHEGGIOAIYGKYAVIETRAFYRGEPYGKVTRFINSGTO
IVOTYNLLQGAKHGEEFFFYPETGKPKLLLMWHEGILMGIVKTWYPGGTLESCKELVNIK KSGLLTTYYPBGQIMATEEYDNDLLIKGEYFRPGDRHPYSKIDRGCGTAVFFSSAGTITK KIPYQDGKPLLN

862415 861801

CT647 hypothetical protein
TTIYIKLLGRLMKKWISILILSFLSLLSILPVLAITINHVKISQRWSDLNSQILTLKVIR
DHEDQVIKHNARISKDRNNLSI IESLNASCKQLRPLSKERERLNKLNSNSLLAQSKEVWER
KRÆLEKSNHOLVWNCEQMHNDFAFVRLEQATEMDNEDIESLFSLFNPENPVAPLVFFTCW
KNTROTTPLGNEVWLTHAEAISRWI

MITGOTPLONE/WLTHAEAISRWI

II

CPC_0766 863785 862394

CTEAG Hypothetical protein

AMPEKLPVYHIGLIKAENNTIKIAILQKTCKGWIVCHCEQIPEGKTWSLPKKYFAAPTTF
STEGSDILVKSSSSSLKNENTIKVALITHLEASLALPWESLIVQPQLCKPTDRGETPLTL
WLAGKNTLKKELSFLSQAGIFPDKLSCRAADIFFLAEQSPLKSLPAYLLIVGGSEEVTCI
FVERHAIAVARSFSNHSTKKSCDDIHATLQYIQETFPQTVLPAIHVAQISPNLQKILEQK
SLELVVCQSMTYGVEDEDWEIYGDTIAAAHHGASRRPLTFPYDATSVSPAAQKHWLLRS
STELTGKYALMATVVVSLGSVLKLKSLSSSASNHFAFACPEEGVLPRSLKAAEKTVKAIGR
KNEASNYPLLPTIPTSEQTLKFLLALGKSSPSIKFSYFSYTMTSYPSKDNPSLPYSALVR
VKÖGGOPEDIPOFLKKISSHPKLOHVSESLEDORSFKLOFTLSS vkoogopedipoflkkisshpklohvsesledorsfkloftlss

CPh_0767

CPR-0767

S63878

864177

CT545 hypothetical protein

NIMLSYLLRTAINVYSFLILAYIPASWYPDCOSARWYQLVSKCVDPFLNFFRRFVPRI

IDPSPFVGLLCLGILPFVILRVLRFIILNIFHSPWLLQYL

CPTE-0768 864144 865163
yohl/nir3-predicted oxidoreductase
yfs55mAaplfirNilLrssivyaplagfsdvpyrcmsalyopglmfcemvkvebilyap
ertsklldynenmrpigaglcgsnpetsgeaakilegigfbliblnccpytbkfrkdsg
sgllktpetlgrildrinsivsipytvxirsgdmehinvebtyniirdagafayfyngr
TRAGGYHGPSKQEYISRAKAAAGKEFPVFGNgDIFSPEAAQAMLTTGCDGVLVARGTLGA
PWTGKQIQDYLTTGSYEKIPFIRRKAAFLEHMRLVEDYYQSETKFLSETRKLCGHYLISA
AKUBFIPSSI MAATSGEIVOU UNDVERDDESSI ETEUKC AKYRFLRSSLAKATSYQEVYQLVNDYEEADDSSLETFVKC

CPI_0769 867763 865121

CPI_0769 867763 865121

SIQOPHAIRLIMKKSLIIVESPAKIKTLOKLLGSEFVFASSIGHIVDJPAKEFGIDVDHDF
EPOYOVLPDKOEVINHIRKLAKCEKVYLSPDPDREGEAIAWHIANOLPDSPLIQRVSFN
AITKNAVTEALKHPRTIDMALVNAQQARRLLDRIVGYKISPILSRKLQRSGISAGRVQS
VALKLVVDREKAIDAFVPVEYWNLRVLMODPKTTKTFWAHLYAVYCKKWEKEIPEGKTEN
DVLLINSEEKARHYAELLEKSSYITIRVEAKAKRRFAPPFITISTLQEASRHFRFSASR
TMSIAQTUYEGVULDSEDSTGLITYMRTDSVRVDPEALTTVRRYIQOTFGKEYLPEKANV
YTTKKMTQDAHEAIRPTDINLTPDKLKNKLSDDQFKVYNLIYKRFVASQITPAIYDTLAV
QITTOTEIDLRASGSLLKFKGFLAVYEEKQDDENDQEEDHPLPPLHAQDALIKEEVSQEQ
AFTKELPRFTEASLVKELEKSGIGRPSTYATIMNIJOSRBYTKENDGREPTLGKIISF
FLETNIFPRIMDIGFTALMEDELELIADNKKPWKLLLQEFYTTFLPVVITAEKEAVIPRIL
TNIECSKCHKGKLVXIWSKNSYFYGCSEYPECDYRTSEFLAFNKEDYAEDTPWDSPCPL
CGGYMKVHGRYCTFLGCEKYPECRGTISIHKKGEIFOEPEIPCPAIGCNGKIFKKRSR
YNKIFYSCGEYPECSVIGNSIDAVITKYSGTEKIPYKKTPTKKKSSAKTTKAAKTPSKK
GKAKSSVKKSSEKKTGPLFLPSPDLAKMIGNEPVSSGEATKKIWDYIKEHQLQAPENKKL
LVPDNNLATIIGPNPIDMFQLSKHLSQHLTKVSNDESSASS

868322 869131

CT642 hypothetical protein
KPRTRNVEKLEFVTSLGSPDDDLITFNKOGL/AGPEEEKVAFLVRSNAMLDAGPETPASF PECLIREOFO I PPEYVEVLY SNEGLOW WEAKETWILLINEST I QURKHHRKAS RWLOMYSRD
EYLAHEAVHAVRMKFHEPVFEEVLAYOTS BWOWRFFOPLFRSPGESYLLLFFTILGUGI
SLWYPAGILIMLVULTMYFLMRLOMAQSYLYRAMKKIPKMLOV PPUWVLLRLTDKEIKMFA FEI-TPVLEHYARKRKLENVRWKQTYQGYFV

870511

CPD_0771
S70511 S00144
Epoll RIA Folymetage Sigma/54
IFYGNSKRLYDSSALDMFQGKOKDZILKYLPSILMOGGLOMLQSPLTELSSYV/9EIIDNP
FFDJSSGEREEMSPYRPINSTRSYLNOTYSPESLYTRLPGIERASTAERRIAHGI
AYMIGDBGLILMHEDFAQELEUTUREK HIKWOTT UDILSPESLASTAGERRIAHGI
AYMIGDBGLILMHEDFAQELEUTUREK HIKWOTT UDILSPESLASFAGSYMMKLERNSS
HUGAYSTYPDGYFLMTRIEFADIMKKFSLSLSELRHILKKALGSTPWGPAAACTVKPMVS
TFLHDTYLFYSSGSWEHEVSTRGLESIKLIKKETFHFYEHLPKEEOKNLSOOILSAKWLIK

NLRKREOTLLOVMETLLPK CK IPAPYPLS I KOLAEDLIGHEST IFRA I ENKAVA APIGIPPLKHLPPRGIHODS ANVLOWIRCH ITATECTPLSDGV ISDRITAKGIPCAR RTVAKYRAQLKILPANKPKKLFYIRSGNSHFRORGF NLRKREQTLLQVMETLLPK APIGIFPLKHLFPRGIHQDS

CPn_0772 872400 870465

uvtD-DNA Helicase

KLGLIMTCISELNEAGRKAVTAPLNPVLVLAGAGAGKTRVVTYRILHLINGGIAPREILA

VTFTNKAARELKER!/NGCASTNEFDVEJVCTFHSLGVFILRRSINLLARENNFTIYDQS

DAEPLIRIAL, JHRIFTGLAFFIOAFF JOAN SCARFELFFIEDLER HTVLFFFITHAGELINGHE

TOTAL SENTENCE SE RVFAUDPUD: 15ARSANIARIAN STOREEADPVAAEILOLIRVCNIKLEDICIFYRTISOS RTFEDALLRRIPYEIIGGLSFYKÄKEIQDILAFLRIFISKSDIVAFDRTVNLPKRGIGS TTIFALTOYAIAQOLPILKACQOÁLDTKDVKLSKKQQEGLQEYLALFPQIEHANTISLR DFIESVVRITGYLEILKEDADTYKDRKSNLEELYHKALESEQONPKTHLELFLDDLALKG SDDDLNITADRVNLMTLHNGKGLEFRVSFLVGLEEQLLPHANSLGGTYENIEEERRLCYV GITRAQDLLYLTAAQVRSLWGTVRMKPSRFLKEIPKDYMIQVR

CPn_0773 8/2485 873195
ung-Uracil DNA Glycosylase
FMONATIDOLPSYMOPOLPLCWREQLKEEWSKPYMOQLLIFLKQEYKEHTVYPEENCVFS
ALRSTPFDQVRVVILG9DPYCKGQAHGLSFSVPEGQRLPPSLINIFRELKTDLGIENHK
GCLQSWANQGILLLVYVLTVRAGEPFSHAGKGWELFTDAIVTKLIQERTHIIFVLWGAAA
RKKCELLFNSKHQHAVLSSPHPSPLAAHRGFFGCSHFSKINYLLNKLNKPMINWKLP

CPn 0774 873183 873425

CTION 1 hypothetical protein LEAPMNEGIHS/CFOKTFRLTAKSVVSMEMLLTTQQLPSAEGMPSVANLEADFLRAEALL AEMREIRGCLEGSLRTLVPSE

CPn_0775 874040 873414
yggV famfly
ERFMKIV/ASSHGYKIRETKTFLKRLGDFDIFSLSDFPDYKLPQEQGDSITANALTKGIH
AANHLGWVIADDTHLRVPALNGLRGPLSANFAGVGAYDKDHRKKLLDLMSSLESLVDRS
AYFECTVVLVSPNGEIFKTYGICEGYISHQEKGSSGFGYDPIFVKYDYKQTFAELSEDVK NOVSWRAKALOKLAPHLQSLFEKHLLTRD

_ 0776 874180 875487

CPM_0776

874180

87-2480

CT605 hypothetical protein

FIFVLKNFYDCLLMFFQFLSFTMKKIFYSFVLLSCIFPYVGCAQVFVGLDRIFSDGEYTR

£IGKKIALISHSAAINSRGQDALSVFYSRKDDCTVELICTLEHGYYGATFTETVGNOPS

RYPNLRSVSLYGVKEYPKEVAEHCDYFYDVQDIGVRSYSFVTVLMQTVKASERYKGQLI

VLDRPMPMGGRIVDGPLNPTTSGSLAIPYCYGMTPGELALFFKKTYAPNANVVVI PHKG

WNRSMTFDETGLIWMPTSPGMPDPQSPFFYAATGILGALSVASIGVGYTLPFKVLGAPMM

CENTROLET MINUT GCULFLIPFFYEDFFGKYKMEMCSGVLLVLQDPKIFFYPETQCTIWG DGEKVADELNRMKLPGVLFLPFFYEPFFGKYKMEMCSGVLLVLQDPKIFYPVETQCTIWG VLKALYPKQVEQTLKSIERIPARRSSICNLFGGDEFLSISHKERYIVWPLRRLCKESRES **FHOLRSSCLLSEYAES

CPn_0777 875586 877178
groEL_2-heat shock protein-60
TSEDRYWWFKSQFEGLSALKRGVHALTKAVTPAFGPRGYNVVIKKGKAPIVLTKNGIRI
AKEIILQDAFESLGVKLAKEALLKVYEGTGDGSTTALVVIDALFTQGLKGIAAGLDPQEI
KAGILLSVEMVYQQLQRQAIELQSPKDVLHVAMVAANHDVTLGTVVATVISQADLKGVFS SKOSGISKTRGLOKRVKSGYLSPYFVTRPETMDVVMEEALVLILSHSLVSLSEELIRYLE
LISEONTHPLVIIAEDFDONVLRTLILNKLRNGLPVCAVKAPGSRELRQVVLEDLAILTG
ATLIGGESENCEIPVSLDVLGRVKQVMITKETFTFLEGGGDAEIIQARKQELCLAIARST
SESECQELEERLAIFIGSIPQVQITADTDTEQRERQFQLESALRATKAAMKGGIVPGGGV
AFLRAAHAIEVPANLSSGMTEGFETLLQAVKTPLKVLAQNCGRSSEEVIHTILSHENPRF
GYNGMTDTFEDLVDAGICDPLIVTTSSLKCAVSVSCLLLTSSFFISSRTKT

CPn_0778 877400 878092
tsa/ahpC-Thio-specific Antioxidant (TSA) Peroxidase
APVAgSDRVEGYEPGGQRESSLVRNNRRVEEEVFMTLSLVGKEAPDFVAQAVVNGETCT
VSLKDYLGKYVVLFFYPKDFTYVCPTELHAFQDALGEFHTRGAEVICCSVDDIATHQQWL
ATKKKQGGIEGITYPLLSDEDKVISRSYHVLKPEEELSFRGVFLIDKGGIRHLVVNDLP LGRSIEEELRTLDALIFFETNGLVCPANWHEGERAMAPNEEGLQNYFGTID

CPn 0779 878502 878095

878502 878095
CT602 hypothetical protein
RFDLIFGMKFTVALFGEAKGSYDTAYFCRSLVDLHNYLGDVSSPGITLAIKTLLSDYNV
VYFRVREEGYCVDSYFFGLHFLNTOTTLKNIIAIGLPGVGNOHIIEASRSLCQKHNSLLL
FFDHDLYDLLTFNQPF

CPn_0780 879241 878591
papQ/amiB-N-Acetylmuramoyl-L-Ala Amidase
HGNKIAVOSLRFMHAKLSFFILLSLLFSGIDCSFLHAAGRSPSLQOVLAEIEDISAKLAS
HEVEIVMLSERLDEQDSKCQKWTAAKPETLAQKIRELESLQKALAKTLAVLTTSVKDLQT
NLOSKLOEIQKDHRALAQDLRLVRRSLLALVDSSSPGAYADFSDPVPENIYIVREGDSLS
KIAKKYKLSVTELKKINKLDSDAIYAGQRLCLQPNKQ

CPn_0781 879851 879198
pal-Peptidoglycan-Associated Lipoprotein
ONCYRSRRKTYPLLGCFPSATDKENTMNIHSLWFLCTLLALLALPACSLSPNYGWEDSCN
TCHHTRRKPSSFGFY-PLYTEEDFPNFTFGET/DSKEEK/YKSSGVAAFRNITFATDSYT
IKGEENLAILTNLVHYMKKNPRATLYIEGHTDEPGAASYILALGARRANAIKEHLRKQGI

SADRLSTISYGKEHPLNSGHNELAWQQNRRTEFKIHAR

CPr. 0782 881077 279777

CHI_0182

TO IF you've accharide transporter

CD IT/MEROLCFOVFFFC FASILVYABELEY/VRSEHITLE IEV/COTOTKOPK [OKYLISCL
TELFF(KD IALIDKLA)FTAASKESSSPLA ISLREHVPQLS/V/LLQSCKTFQYILCFFT HON
LEVOP/VRTHHAADTVINALTGI FG ISAGK IVFALSSLGY DOKLKOGELWTFDYD XKNLAP THE ACTION AND ALTER ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION ACTIO

COMMON TO THE PROPERTY OF THE

RELIBER 881109 8 hypothetical protein

EMMKAFBA LYTLYC, HICKLITTT NEVT DUBKKBTÖBKYŁĞEKTALIĞEKBIALLEGAAADE

AKT TRP::VATQPQKQAKC3PPQENVQKALQKPTPKVTH SPAPTVAKKTTATEKP WPSTAQLTMHSELKAT PRSTTKKNTQLGKTQLQTLSEVAQALSLHVDK1EKSET QEDELCELFRTH1ALPSKGYVR1KLVLGPNGE1QECSFL SAADKOLLTORIOALPFQ KFLEKYKVSKNISFHIKLVSNES

982359 881892 exbD-Biopolymer Transport Protein
DRADGIFTTEFYPGOIPOYKIMKYPFTEEIEEEPLVNLTPLIDIVFVILMAFIVAVPLIK CONTALAGOROS, AMENORADO INVERMENDADA DE COMERCIO DE DESENTAR EN LA HIRAYA TPELLODGETGERTYON/KINATEAN/SEHELIIVALON

883039 CPL_0785 883039 882296
exbB/to1Q-polysaccharide transporter
DHLYFETLSVNKDFYSMVHFSHNPIIQAYTEADFFGKSIFFCLLILSVCTWTVLHQKLAI
OKNFLKAGKSLKDFLIKNRHAPLSLDIHPELSPFADLYFTIKRGTLELLDKNRQSAPDRG
PILSSEDIOSLETLLGAIHPKYKALLHKNSFIPATTISLAPPLGLLGTVWGILVAFTHIS
SGSSGNSAIMEGLATALGTTIIGLFVAIPSLIAFNYLKAHSSELISEIEQTAYLLLNSIE CPn_0785 VKYRNTNI.

CPn_0786 883137 885293
dsbD/xpra-Thio:disulfide Interchange Protein
NHGVILNKFKTYLOTALIAPFFSFPALSGSFSSIOAEEITOQVNNPGAELLSEGSYIPGL
OTFRLGIKITASKGSHIYWKNPGGIGSPLKISWGJPKGFVVEEEHWPTPKVFEEEGTTFF
GYEDSALIVADVRAPEGYTFGGEVELRAQVEWLACGDSCLPGNVDLKLTLPYEEKEPSLY
PDTHAEFTKTLHAQPRVLENDHSVQVAQGKGMEIILNISKKINATKAWFVSEKADKLFAY
AETSYSGGTGTAWRLKVKNLSGVQKNEKLHGILLADHTGREVESLTHSEVLGQTGSAV
AGLSQYITILIMAFLGGVLLNIMPCVLPLVTLKVVGLIKSAGEHRSSVIANGLWFTLGVV
CCFWGLAGVAFILKVLGHNIGWGFQLQEPMFVATLIIVFFFLASIGLFFMGTMFANLG
GKLQSSEMKSSNNKAVGAFFNGILATLVTTPCTGPFLGSVLGLVMSLSFLQQLLIFTAIG
LGMASPYLVFSVFFKMLSVLPKPGGMMSTFKQLTGFMLLVTVTVMLVWIFGSETSTTSVV
LLGGIMLAGIGAMILGRWGTPVSPKNGVCASLLFFAFLGGGISVSGLASTYFAEPQGTV
SVNEDSLWQPFSLEKLAQLRAQGRPVFVNFTAKWCLTCQMWKPVLYGDAVQKWFETHGIV
TLEADWTRKDPGITEELARLGRASVPSYVYYPGDNSAPVVLPEKITQNLLEDVVSRFVR

CPn_0787 885604 986401

yabD/ycfh-Php superfamily (urease/pyrimidinase) hydrolase

TRROPVDLADAHVHLSDDAFEEDINSVLQRAQDSGVSLVWNVTTTEKELNRSFAYAERFP

KIRFCHVGGTPPQDVDQDIEEDYRNFHAAAHSKKLAAIGEVGLDVCFATEEGIARQKEVL

QRYLALSLECELPLUVHCRGAFNDFFRHLDQYYRNDPRSRPGMLHCFTGTLEEAQELISR

GWFISISGIVTFKNAQDLRDLVVELPLEHLLIETDAPFLAPVPYRGKKNEPAHVLHTINA VANAKGMFPQELAALAYKNVLRFLHG

887432 886521 LHTED/AB 88521 887432
sdir-Succinate Dehydrogenase
SLWSLEMSRHEICPEVSHKKCKYYSTFIFRCIHSLAGIAFTFFLCEHLFTNMLASSYFS
OGKGFVAMANGFHKIPGLKIIEVAGLVLPFLCHAIIGIVYLFOGKSNCYSGDSRPHLRY
AKMASYTWORWTAWILLFGIAFHVYHLHFIRYPVHVDHGFTTYVAVDIOPSTYDVINGT
KGFLTLNLPNTEASSIEVSRHDLGGADAALLSERNSYLLTPSAGTAFLYVVRDALGSLFI
ALÄATILVIAAAFHGFNGLWTFCCRWGVVVSLRWGGVLRIVCYLAMIVVFFMGVSAVWNL
YSVA YSVA

CPt-0789 887436 889316

Succinate Dehydrogenase

OMDENRKVIVVGGLAGLSAAMQLANLGIIVELVSLTKVKRSHSVCAQGGINAALNILKPE

EEDSBYVHAYDTIKGGDFLADDPFVLEMCLAAPRIIN-LIDNEGCPFNRGPSGNLDVRRFG

GTLYHRTVFGGASTGQLHYTLDEGOVRREHAGRVIKRENHEFVRLVTDHSGRACGILM

NLFNNRLEILRGDAVIIATGGPGVIFKMSTNSTFCTGAANGRLFLOGMAYANPEFIGHP

TAPTGRKKLRLISESVRGEGGRWWPGDSSKRIVPPDGSERRCGETGAPWYFLEDMFPAY

GNLYSRDVGARAILRVCLAGLGIDGRMEAYLDVTHLPEKTRIKLEVVLDIYKKFTEDDPN

TVFHRIFPAVHYSMCGAWVDWPAADDPDDSRFROMTNIPCFNGGESBFOYHGANLGA

NSLLSCLFAGLVSGDEASRFIEAFGASQATSSDFDRALOCKEENARLLSASCKENIFVL

HEPTAK HWRNVTVKRNNRDLQETHØKLKFFRERLKNVSVLDSSFFANKSFHVROMGPM

LELALAITKGALLRNEFRGSHYKPEFPERDDEHWLKTTVAVYAPEEPEISYLPVDTRHVA

PTERDYTKSSTGKIELTNIPDNIRLPI

889279 890103 CYNEW 790 889279 89103
SAMBS SUCCINATE DENYTOOGRASE
NSRIFLIISVYPYRKREMMENLETFILKIYRGVPGKQYWESFELPLKPGENVISALMEIE
KREVNILGEKVNPVVWEQGCLEEVCGSCSILVNGVPRQACTALIGYYIDATOSREIVLAP
LTKFPLIRDLIVDRSIMFONLERIGGWVAADIEGETFGPQVTGEGCELYALSQCMTGGC
CTEACPQIDNKSDFIGPAAISQARYFNTYPGDKRSKKRWRALMSKGGIEGCGQAHNCVRV
CPKKLPLTESISAVGREISKFSLRSLFSALFKKKK

CPn_0791 893104 890111
CT590 hypothetical protein
CLRSSKRIVVEDISDRNMYSCYSKGISHNYLLHPMSKLDIFVFDSLIANQDONLLEEIF
CSEDTVLFKAYRTTALQSPLAAKNLNIARKVANYILADNGEIDTVKLVEAIHHLSQCTYP
LGPHRHNEAODREHLLKMLKALKENFKLKESIKTLFVFSYSTIONLIHTLALNPOTILS
TIHVRQAALTALFTYLROPOVGSCFATAPAILHTGYPERFLKDLNDLISSGKLSRIVNQR
EIAVPINLSGCIGELFKPLRILDLYPDPLVKLSSSFGLKKAFSAANLIETLGDSEAOIQ
LLSHQYLMGKLONVHETLTANDIIKSTLLHYVOLQESTVRAIFFKEGLFSKEQVAFSTQH
PRELSEIGRYYHYLHAYEEAKSAFIHDTONPLLKAWEYTLATLADASQPTISNHIRLALG
WKSELPHSLVSLVTHFVEEEVENIRILVQQFEQTYHEARSQLEYIEGRMRNPLNNQDSQI
LTMDHMRFQDELNKALYEWDSAQEKAKFFHLDFELLSFYTKJFLYBVAYSVAFIGFFA
LLYANAPAGFRILFTHGRTHPNTWSPIYJINEFIRFLSEFFTSTESELLGKHAVINLEKE
TSRLVHNITAMLHTDVFQEALLTRILEMYQLEVPPSILNHLDQLSQTPWVVVSGTVDTL
LLDYFESESEPLTITEKHPENPHELAAFYADALKDLPTGIKSYLEEGSHSLLSSSPTHVFS
IIAUSPLFREAWDNDWYSYTWLRDVWYKGHOPFLGTILPQLSIYAFIENFCNKYALOHV
VHDFHDFGSDHSLTLPELYDKGSRFLSSLFTTDKTVALIYIRRLLYLLAVREVPVVSGOQL
PEVLDNYGGYLGISSRITYEKFRSYLEETIFKMTLLSSADLRHIYKGLLMQSYQKIYTEE
LTYLRLTTAARHHINLAYPAELLFADSNMPSIYFGFILHPGTTEIDLMKFNYAGLQGQPLD
HIGGELFATSHPWTLYANFIDYCMPPPPGYRSRLPKEFF 893104 890111 CPn_0791

893108 CHILDFOR 844/55 89310H
CTTSB9 hypothetical protein
RHHLINIKGIRIMKTETKRYLEFFELVIPIBLLHILMVVGFEGEGAAKANLVOVLHTRA
THISTIEFERKLT HKLELDKLANTLALKSYASESAEPYAQAYNEMMAGMTDESUGLIDP
ETCOUVRTYHIGDBY IRYLKOHPEMKKKISAANGKAFLLTTIGKPULHVLILVEDVASWDS
TTTCGLUUGFYEMERGUOLFOSIH HTKCNICLVMYYGEVLEGAQDRESSEVFEDLDFNL
FGFQARSISAIEIEKAYGILGENLITVSINKKRYLGLVIMKIPIQATYTLSLVPVSDLI
QGALKVPLHICFEYVYAFLLMMWIFSKINTKLNKPLGELTFCMEAAWRGNINVRFEPQPY GYEFNELGHIFNOTTLLLLLI

GYEFNELGHIFNOTTLLLLLI

DIDYHOGEKLÖREZGA TÜÜÜLÜGALLÜPDFFFFFKV

DTLLGIIGLAGDYGLEBIYLATALBARILELAYASSD

VSLOKISKOTADSFOKTTEGIEAVVAMTFIKYVEKDRÜLELLÜLSEGAFTNELORGESFV

RLELEHNOALOPODRLICLTYGEDILKYFSOLPIEELKBPLNELMTENLIDSLTMHLNN ETEHSADGTLTILSF3

996838 89491 rbsU-sigma regulatory family protein-PP2C phosphatase (RsbW

697123 898004 CPn 0794

CPn_0795 898008 899195
NO FODUST NOMOTOG PRESENT IN GENEBARK/EMBL AS OF 11/7/98
GTLGGANSSATGVSSDGSVIVGQAQTADKSVHAFQYYNGEMKDLGTLGGTSSTAKTVSPD
GKVIMGRØGIADGSWHAFMCHTDFSSNIVLFDLDNTYKTLRENGRQLNSIFNLONMALQR
ASDHEFYEFGRSNIALGAGLYVNALQNLPSNLAAQYFGIAYKIRPKYRLGVFLDHNFSSH
VPNNFJVSHNRLWMGAFIGWQDSDALGSSVKVSFGYGKQKATITREOLENTEAGSGESHF
EGVAADIEGRYGKSLGGHVRVOPFLGLOFVHITRKEYTENAVQFPVHYDPIDYSTGVVYL
GIGSHIALVDSLHVOTRMGHEONFAAHTDRFSGSIASIGNFVFEKLDVTHTRAFAEMRVN
YELPYLQSLNLILRVNQQPLQGVMGFSSDLRYALGF 898008 899195

CPn_0796 899280 901340
No robust homolog present in Genebank/EMBL as of 11/7/98
SELYSSYLOPCLNMSIVRNSALPLPCLSRSETFKKVRSHMKFMKVLTPWIYRKDLWYTAF
LLTAIPGSFAHTLVDIAGEPRHAAQATGVSGDGKTVIGHKVPDDPFAITVGFGYIDGHLQ
PLEAVRPQCSVYPNGITPDGTVIVGTNYAIGHGSVAVKWNGKVSELPMLPDTLDSVASV
SADGRVIGGRNINLGASVAVKWEDDVITQLPSLPDANNACVNGISSDGSIIVGTMVDV
SMRNTAVQWIGDQLSVIGTLGGTTSVASAISTDGTVIVGGSENADSGTHAYAYKNGVMSD
IGTLGGFYSLAHAVSSDGSVIVGVSTNSEHRYHAFQYADQMPULGTLGGPESYAGVSG
DGKVIVGRAQVPSGDWHAFLCPFGAPSFAPHGGSTVVTSQNPRGMVDINATYSSLKNSQ
OQLCRLLIGHSAKVESVSSGAPSFTSVKGAISKQSPAVQNDVQKGTFLSYRSQVHGNVQN
QQLLTGAFMDWKLASAPKCGFKVALHYGSQDALVERAALPYTEQGGGSSVLGGFGGVQV
RYDFNIGETVVLQPFMGIQVLHLSREGYSEKNVRFPVSYDSVAYSAATSFMGAHVFASLS
PKMSTAATLGVERDLNSHIDEFKGSVSAMGNFVLENSTVSVLRPFASLAMYYDVRQQQLV
TLSVWMNQQPLTGTLSLVSQSSYNLSF

CPn_0797 901552 902694

No robust homolog present in Genebank/EMBL as of 11/7/98

VLILTWINVLTKLGLMMSKKIKVLGHLTLCTLFRGVLCAAALSNIGYASTSQESPYQKSI
EDWKGYFTPTOLELLSKEGMSEAHAVSGNGSRIVGASGAGGGSVTAVIWESHLIKHLGTLG
GEASSAEGISKDGEVVVGMSDTREGYTHAFVFDGRDMKDLGTLGATYSVARGVSGDGSII
VGVSATARGEDYGWVGVKWEKGKIKQLKLLPQGLWSEANAISEDGTVIVGRGEISRNHI
VAVKWNKNAVYSLGTLGGSVASAEAISANGKVIVGWSTTNNGETHAFFHKEETHHDLGTL
GGGFSVATGVSADGRAIVGFSAVKTGEIHAFYYAEGEMEDLTTLGGEEARVFDISSEGND
LIGSIKTDAGAERAVIFHHKK IIGSIKTDAGAERAYLFHIHK

CPn_0798 902810 903856
No robust homolog present in Genebank/EMBL as of 11/7/98
VVFEIIFVVRVPMKKTCCQNYRSIGVVFSVVLFVLTTQTLFAGHFIDIGTSGLYSWARGV
SGOGRVVVGYEGGNAFKYYDGEKFLLEGLVPRSEALVFKASYDGSVIIGISDQDPSCRAV
KWYNGALVDLGIFSEGMQSFAEGVSSDGKTIVGCLYSDDTETNFAVKWDETGMVVLPNLP
EDRHSCAWDASEDGSVIVQDAMGSEEIAKAVYWKDGEHLLSNI FGAKRSSAHAVSKDGS
FIVGEFISEENEVHAFVYHNGVIKDIGTLGGDYSVATGVSROGKVIVGHSTRTDGEYRAF KYVDGRMIDLGTLGGSASFAFGVSDDGKTIVGKFETELGECHAFIYLDD

905001 903940 No robust homolog present in Genebank/EMBL as of 11/7/98
KREENMAAIKOILRSMLSOSSLWMVLFSLYSLSGYCYVITDKPEDDFHSSSAVKWDHWGK KREEMMALIKQILKSMLSQSSLMMVLPSLYSLSGYCYYITUKPEDDFHSSSAVMKKSKATG TTLSRLSNKKASAKAVSGYGATTVOFIKDTWSPTYAVRKNYMCTKELPTSSAVKKKSKATG ISSDGSIIAGIVENELSQSFAVTWKNNEMYLLPSTWAVÇSKAYGISSDGSVIVGSAKDAW SRTFAVKWTGHEAQVLPVGWAVKSVANSVSANGSIIVGSVYDASGILLAVKWEGOTTITHL GTLQCYSAIKAVSNNGKVIVGRSETTYYGEVHAFCHKNGYMSDLGTLQGSYSAAKGVSAT GKVIVGMSTTANGKLHAFKYVGGRMIDLGEYSWKEACAHAVSIDGEIIVGVQSE

906550 905249 eno-Enolase eno-Enolase
RKEIKIMFEAVIADIQAREILDSRGYPTLHVK/TTSTGC/GEARVPSGASTGKKEALEFR
RKEIKIMFEAVIADIQAREILDSRGYPTLHVK/TTSTGC/GEARVPSGASTGKKEALEFR
DTDCPRYQCKGVLQAVKNYKEILFELVKGCSV/EGSLILGSLMMDSDGSPNKETLGANAIL
GVOLATAHAAAATLRRPLYRYLGGCFACSLPCPMMNLIN/ZMHADNGLEFQEFMIR PIGA
SSIKEAVNMGADVFHTLKKLLHERGLSTGVGDE/GFAPHILASNEEALELLLLAIEKAGFT
RGKDIGLALDXAASSPYNKTGTYDGSRYEEGIALLGHLCDRYPIGGIEDGLAFEDYDGW
ALDTEVLDEKVGIVGDDLFVTNPELILEGISNGLAN:ZVLIKPNQIGTI/TETVYAIKLAQM
AGYTTIIGHRGGTTYPTIADLAVAFNAGQIKTGGLSGERVAKYNRIMEIEEELGSEAI PTDOMVESYEDSEE

908709 906727 CPI_9801 908/09 906727

OVER EXTRICTORS ARC SUBURIT. B
ITEMMEDIATE ARC SUBURIT. B
ITEMMEDIATE ARC SUBURIT. B
ITEMMEDIATE ARCHOMER
SGH IVQQI IRPTG IPDP GFLSELE IPAAYLHSG I VEYOFRUPSAFONRPLTYEEAOKYFRKZIYVSATPGD1 WPETERATION DOWN POTTER OF THE WASHINGTON OF THE

910306 909752 CT584 hypothetical protein FMAAKTKTLELEDNVFLLLEGNLKRIFATPIGYTTFREFQNVVFNCANGQQEIANFFFEM LINGKLTQELAPQQKQAAHSLIAEFMMPIRVAKDIHERGEFINFITSDMLTQQERCIFLN

RLARVDGQEFLLMTDVQNTCHLIRHLLARLLEAQKNPVGEKNLQEIQEEITSLKNHFDEL

CPn_0804 911074 910310
gp60-CHLTR Plasmid Paralog
EIFSSMGNLKTLLESRFKKNTPTKMEALARKRMEGDPSPLAVRLSNPTLSSKEKEQLRHL
LQHYNFREQIEEFDLTQLCTLSAEVKQIHHQSVLLHGERITKVRDLLKSYREGAFSSWLL LTYGNROTPYNFLVYYELFTLLPEPLKIEMEKMPRQAVYTLASROGPQEKKEEIIRNYRG ERKSELLDRIRKEFPLVETDCRKTSPVKQALAMLTKGSQILTKCTSLSSDEQIILEKLIK KLEKVKSNLFPDTKV

CPn 0805 911846 911067 minD-chromosome partitioning ATPase—CHLTR plasmid protein GP5D GYASRMKTIAVNSFKGGTAKTSTTLHLGAALAQYHQARVLLIDFDAQANLTSGLGLDPDC YDSLAVVLQGEKEIQEVIRPIQDTQLDLIPADTWLERIEVSGNLAADRYSHERLKYVLGS VQDKYDYVIIDTPPSLCWLTESALIAADYALICATPEFYSVKGLERLAGFIQGISARHPL TILGVALSFWNCRGKNNSAFAELIHKTFPGKLLNTKIRRDITVSEAAIHGKPVFATSPSA RASEDYFNLTKELLILLRDI

CPD_0806 913816 911867

LHTS-Threonyl LRNA Synthetase

ANNESPPNMEAWNKMIOVTCOKNYEVLEGTTAAELAKQLKNSHQFIGVLINERPRDLS

THLWEGDTLVFLITSEDPEGREIFLHTSAHLLAQOVLRLWPDAIFTIGFVIDHGFYYDFAN

LSISESDFPLIEDTVKQIVDEKLAISRFTYGDKQQALAQFPQNPFKTELIRELPENEEIS

AYSGGEFFDLCRGPHLPSTAHVKAFKVLRTSAAYWRGDPSRESLVRIYGTSFPTSKELRA

HLEQIEEAKKRDHRVLGAKLDLFSQQESSFGMPFFHPRGMIVWDALIRYWKQLHTAAGYK

ELETPOLMRQLWEVSGHMDNYKAMNYTLQIDDEDVAIKPMOPSCHLYYKTRLHSYKEF

PLRVAEVGHVHRQEASGALSGLMRVRAFHQDDAHVFLTPEOVEEFILNILQLVSTLYGFF

GLEYPLLEISTRPEKDTIGDDSLWELATDALNRALVOSGTPFIVRGEGAFYGFKIDHVW

DAIGRTWOGGTIQLUMFLPEFFELEYTTAGGTKSVPVMLHRALFGSIERFLGILIENFKE

RFPELLSFEQURIITVADRHIPRAKELEEAWKRLGLVVTLDDSSESVSKKIRNAQNMOVN

YNTIFLIGDHEINENVLAVRTRDNRVINDVSVERFLNTILEEKNSLSLTALL 913816 911867

CETTO807 913950 914879
CTT50 hypothetical protein
TLOTGLHMSLFLVFLTAFIWSSSFALSKLVMNASAPIFATGARMVIAGAILALAAWFRGS
FVGISKKIFLYIVLLALTGFYLTNIFEFIGLOSLSSSKTCFIYGLSPLMSALFSVIQUE
KYFEKKVIGLSLGLVSYICYLTFGGGDDSQPHTWOIGLPELIIGAASLASFGWTLKRO
IEKQSTLSVTAINAYAMLIAGMLSIMHSAVVEFWRPLPVQDISQFLYATLALWISMLIC
YNLTAKLLRKYSSTFLSFCNLVMPLYSGFYGWILLGEKGVSLGLVLAVAFMVAGCELIYH EEFROGYIVS

CPALOBOB 916398 914956
CT579 hypothetical protein
LKKEPSWALKSLKKRMPOSAEPSLAHIKPIIFKGACIAMTSGVSGSSSODPTLAAQLAOSS
CKAGNAQSGHDTKNVTKQGAQAEVAAGGFEDLIQDASAQSTGKKEATSSTTKSSKGEKSE
KSGKSKSTSVASASETATAQAVQGPKGLRQNNYDSPSLPTPEAQTINGIVLKKGMGTLA
LLGTWNTLMANAAGESWKASFOSONQAIRSQVESAPAIGEAIKRQANHQASATEAQAKQS
LISGIVNIVGEPTVSVGAGIFSAAKGATSALKSASFAKETGASAAGGAASKALTSASSVQ
QTMASTAKAATTAASSAGSAATKAAANLTDDMAAAASKHASDGASKASGGLFGEVLNKPN
WSEKVSRGMNVVKTQGARVASFAGRALSSSMOMSQLMHGLTAAVEGLSAGQTGIEVAHHQ
RLAGQAEAQAEVLKQMSSVYGQQAGQAGQLQEQAMQSFNTALQTLQNIADSQTQTTSAIF
N

CPn_0809 917794 916307
CT578 hypothetical protein
DTNMSISSSSGPDNOKNIMSQVLTSTPQGVPQODKLSGNETKØIQOTRQKNTEMESDAT
IAGASGKDKTSSTTKTETAPÇQGVAAGKESSESQKAGADTGSYSGAAATTASNTARKITARG
TSIEEASKSMESTLESLOSLSAAQMKEVEAVVVAALSGKSSGSAKLETPELPRKGVTPARG
EVIEIGLALAKAIQTLGEATKSALSNYASTOAQADOTNKLCLEKQAIKIDKEREEYQEMK
AAEOKSKOLECTMDTVITVMIAVSVAITVISIVAAIFTCGAGLAGLAGAAVGAAAAGGA
AGAAAATTVATOITVQAVVQAVKQAVITAVRQAITAAIKAAVKSGIKAFIKTLVKAIAKA
ISKGISKVFAKGTCMIAKNFEKLSKVISSLTSKWVTVQVVVVAAPALCKGIMQMQLSEM
QONVAQFOKEYOKLQAADMISMFTQFWQQASKIASKQTGESNEMTQKATKLGAQILKAY
AAISGAIAGAHKTNNF

918193 CPR_UALD VISUS STARS
CT577 hypothetical protein
GEIMNKKPKKTKKAVOSKAAPVKRVEESQEAAJOQLELAVSDLYKELPLAQTFASLTDK
NO INSTIAALECTLESLHILEELTOGLFFSAGEFANFAKELSSVVHGLKNLTTVVNKOMVK

918933 ∍19**2**08 CPIC_0811 918933 919208

LCHH-LOW CA RESPONSE FROGER FOR STATEMENTESPLAELAAOKKAKADDLEOVIIPVPT

EEETKKALSHIFTASISIKSLDLOOTLOGJOYLLEETYTVAYTEYSQUKYNEAYSLFOLLAA
AOFONYKYMLAISSKYROLIILYNEAAFOFFLAFDAOPONFI PYYY LADSILLKLOOPEESN
NFLOVIMOTOCHNITEKTI KERCOTIKOS LEKOMAGETKKAFTKKPAKSKI YTTINKKSSK

ED

C14c_9812 920862 MULL-DNA MISMALCH R
GILICWLCNLTKAPMSTRRE DPLT NO LAAGEVIENSVOVVKELIEISLDAGADET
EIETLOCOOGALIEDNOCGFRAEDIF VALORHATSKIREFSDIFGLNSFSFRGEALPSI
ASISKMELOSSIECDEGVRTVIHGGDIVSCEPCARGLOTTVIVNSLFYNVPVRRGFOKSM
OSDRIGIRKLIENRILSTANIGWSWISECHHEIQIAKQQJFQERVAVWHOODHFWODALTI
DKEANGVRIVOVLOSPSFHRFTROCOKIFINDRPIESLFISKKVGDAYALLLPLHRYPVF
VLKLYLPSSWCDFWYHPOKIEARIIKEELVGDIKEAIVETLACPPGILGRTHQEIEESD
SVPLFMFRHLETSDVQEEESVEFDONLFAYSSEDVGLEKOFYTSGRGFKSOMDWIYSSDV
FUTCHSIEVZIARDICTERIVATERILFTVARFRIESELSDS, FRIEMNI, JAAGEST ST. INTERIA
FREEDINGVZIARDICTERIVATERILFTVARFRIESERS ST. INTERIA BERGEBART ST. INSEALT
ALMKETLIDATESKHOWEDVSWIKLERSVOKEREGFDIARIRRILLIDATESKHODETMON mut L-DNA Mismatch R ALMKETLTQATFSKHQHVFDV9WLKLLWGVGKPEKGFDJARIRRLILDSDFMEG

CPn_0813 92/843 921934
pepP-Aminopeptidase/P
TLILIKONHMSHDRILRAGRALSEHNLDAILVEKSEDLAYFLHDEAIAGILLIGQEVMF
FVYRMOKDLYSHIGRVPLYFLTODUVADLSLYVQKQRYQKIGFDSASTVYHKFAQRQVLP
CLWEPLECFTEKIRSIKSEEEIRRMQEAAALSAGYDYVLTLLREGITEKEVVRQLRAFW
AEAGAEGPSFPPIIAFÆHSAFPHSIPTDRPLKKGDIVLIDIGVLLNGYCSDMTRMTALG
TPHPKLLESYPVVVEXQRRAMALCKEGVLWGDIDAEAVRVLREHHLDTYFIHGIGHGVGR
HIHEYPCSPRGSQVVLESGMTITVEPGVYFPGIGGIRIEDTLCIDKNKNFSLTARPVISE

CPn_0814 921996 923357
CT814.1 hypothetical protein
FFLFFKLSYNFIFNLPLTMYQLLSIGYSFVSFIALLWMLCYSPNYVTDLYRISLSAEESL
GGIRAFPQAEĞLLGGACALNFPDLEERLPDLRKELLFLGSNDRPDACGGKFSLQLASSKE
CYIAALKERVILNYTNSSRGFVYSFSFRGVPTELMIECFSVSVDGRVEVKVRLQGLHKEL
ISKPRDCETLFLNPPANKLDCWEIAGFRVDASFPVKQKIRRIGVDKFLLMHGGAEYADKA
TKERVDFYSSDEENYSRVLAVGDVLLWGDCNCVTCGFFGASSRAPLFEVKRIDDKVMIA
DLWNVGGTQROTISLVKGVPSPIEINEVIREIEFTGMRSWSKPIVLVGQGRILLSFDDWV
LRTAKGWEKLSRADQIGDYVTGKVTGPLLVFEKLEKDLRGFVLRGHMFNAQRTLVETISL
PLKCGFFEDVAVSOEVSSNTRSAAAHPGATNRGCG PLKQGFEPAVASQEVSSNTRSAAAHPGATNRGGS

CPn_0815 923361 925622 gspD/pilQ-Gen. Secretion Protein D mvFrnSLLHLVALSGHLCCSSGVALTIAEGNASLEHSGRGADDYEGMASFNANMREYSL QVSKLYEBERKLRASGTEDEALWKOLIRRIGEVRGYLREIEELWAAEIREKGGNLEDVAL yNHPETTIYNIVTDYGTEDSIYLIPQEIGAIKIATLSKFVVPKESFEDCLTQILSRLGIG WHEPTTIYNLVTDYGTEDSIYLIPQEIGAIKIATLSKFVVPKESFEDCLTQILSRLGIG
/RQVNSWIKELYMMRKECGSVAGVFSSRKDLEALPETAYIGFVLNSNVDAHTNOHVLKKF
INPETTHYDVIAGRWIFGSAGEVGELLKIYNFVQSESIRQEYRVIPLTKIDROMISIL
NAAFREDLTKDVSEESLGLRVVPLQYQGRSLFLSGTAALVQQALTLIRELEGIENPTDK
TVFWNVKHSDPQELAALLSQVHDVFSGENKASVGAADGCGSQLNASIQIDTTVSSSAKD
GSVKYGNFIADSKTGTLIMVVEKEVLPRIQMLKKLDVPKMYNIEVLLFERKLAHBOKS
GLNLLRLGEEVCKKGCSFSVSWAGGTGILEFLFKGSTGSSIVFGYDLAYOFLMAQEDVRI
NASPSVYTMNQTPARIAVVDEMSIAVSSDKDKAQYNRAQYGIMIKMLPVINVGEEDGKSY
ITLETDITFDTTGKNHDDRPDVTRRNITNKVRIADGETVIIGGLRCKQMSDSHGIPFLG
DIRGIGKLFGMSSTSDSLTEMFVFITPRILEMPVEQQERKEEALLSSRPGEREEYYQALA
ASSEAARRAMKKLEMFPASGVSLSOVERDFYVTZ ASEAAARAAHKKLEMFPASGVSLSQVERQEYDGC

925600 927102 gspE-Gen. Secretion Protein E RGKNTMAASILSQELLDILPYTFLKKHCLLPIEESSEAITIAHATATSVIAQDEVKLLIK RGKNTMAASILSQELLDILPYTFLKKHCLLPIEESSEAITIAHATATSVIAQDEVKLLIK KPVRFVLKEESEILQRLQQLYSNREGNVSDMLLTMKEEDGTTISEEEDLLETTDTIPVR LLNVILKEAIEERASDIHFEPCEDSMRIRYRIDGVLHDRHSPPSHLRSALTTRLKVLAKM DIAEHRLPQDGRIKIHIGGQEVDMRVSTVPVIYGERVVLRILDKRNVILDIAGLHMPKGT EILFKDTITAPEGILLVTGPTGSGKTTTLYSVLQELKGPLTNIMTIEDPPEYKLPGIAQI AVKPKIGLTFARGLRHLRQDPDILMVGEIRQDETAEIAIQAALTGHLVVSTLHTNDSI AIPRLLDMGIESYLLSATLVGVVAQRLVRTICPYCKVAYTPENQEKSFLASLGKDTEMPL YRCQGCVHCFRSGYKGRGGITFELFNTLFRSEVASNRPYHILRETAEQNGFLPILEHGI ALAYSGETTLAFURNVKRCD ALAVSGETTLAEVLRVTKRCD

CPn_0817 927106 928287
gspF-Gen. Secretion Protein F
GGRMPRYRYTYLDPKERKRKGYLEALHIQEAREKLAQENIQVLDIREVALRRMSIKSTEL
IVFTKQLLLLLRSGLPLYESLVSLRDQYHEQRMGLLLTSFMETLRSGGSLSQAMAAHPNI
FDHFYCSGVAAGESVGNLEGCLQNIIVVLEERAQITKKNVGALSYPCVLLVFSFAVMLFF
LLGVIPSLKETFENMEVKGLTKIVFGVSDCLSAYRYLFLGFASALITVGILMRHRIPMKK
ILEKLLFALPGTKKFVVKVAVNRFCSVASALIKGGGTLIEGLDLGCDAIPYDRLKTDMRD
IVQAVIGGGSLSQELAQRSWYRKLAIGMIALGEESGDLADVLGYVAHIYNEDTQKTLASI
TSCHCODUIT ISLGCI LGUJMI ALI DITSNITGUT TSWCQPVILIFLGGLIGVIMLAILIPLTSNIQTL

CPn_0818 928158 928682
predicted OMP [leader (16) peptide]
GYTKNVGFDNVVVSTRDSDFSwWPDRCDHVGNIDPTHKQYPNIIKCVLRGVGMKRQKRKQ
SITLIEMMVVITLIGIIGGALAFNWRGSIHKGKVFQSEQNCAKVYDILMMEYATGGSSLK
EIIAHKETVVEEASWCKEGRKLLKDAWGEDLIVQLNDKGDDLVIFSKRVQSSNKK

CPn_0819 929117 928956 CT558 hypothetical protein ASLYGYCLFLIWEKFHNNIGKANFHLKIITTDFLTDIYIVTIRDPIAYPLTGIC

929042 929659 CPH_0820 929042 929659
CT567 hypotherical protein
DESLPCRCCCCTFFRSETSSIRTEMPMCNSIAMKKOKRGFVLMELLMSFTLIALLLGTLG
FWYRKIYTVOKOKERIYNFYIEESRAYKOLRTLFSMSLSSSYEEPGSLFSLIFDRGVYRD
PKLAGAVRASLHHOTKODELBRICNIKDOSYFETORLLSHVTHVVLSFQRNPDPEKLPE TIALTITREPKAYPPRTLTYQFAVGK

929637 CPI_0821 929637 930668

CT566 hypotherical protein

ITTNERLONKEMOPE FETLICLIFOLVOLVAFDAANARKECACAGTTERGENEEG IKRSACA
ELEVGEKORHAJA LERIOKUNGUVAFDAANARKECACAGTTERGENEEG IKRSACA
ELEVGEKORHAJA LERIOKUNGUKVAFDAANARKECACAGTTERGENEEG IKRSACA
ELEVGEKORHAJA LERIOKUNGUKVAFDAANAR LONKOE LEEPGAQUG
HUNGUNGUKVAFTA LERIOKUNGOODICLIFELITYEEKOLGICKUNLI FINDELLEAVI
DHEPAYRETSIJJRISTI MEAVIKROEHA TOEHKOAAALELEKTRODERLELRIKMOLLLERY
DILIELJIKKMEDYTLIGA DYLELVDELYKA LORCGERKOIKU 1930668

CEN_0822 43044930000 931225 CENSO, hypother (cd.) protein FELTIVLISTIENTS IGERMADETPREISSIKESSSOFTSLIKRKVKDLIENTKVSKWKKFL SHRAGEATG SCEVIAG I FADETSWARD SEFTARSTOVE FINETRKMESHLOSYS FANGE (K

CPn_0823 132424

CPD_0823

YSCT/SpaR-YOPT Tranlocation T
FYALQVRFSKTSINCHKELMGISLFELFSNLGSAYLDYIFQHPPAYVWSVFLLLLARLLP
LFAVAPFLGAKLFPSPIKIGISLSWLAIIFPKVLADTQITNYMDNNLFYVLLVKEMIIGI VIGFVLAFPFYAAOSAGSFITMOOSIOCLEGATSLISIEOTSPHOILYHYFYTIIFMLOO CHRIVISLLLOTLEVIPIHSFFPAEMMSLSAPIWITMIKMCOLCLVMTIQLSAPAALAML BUILERIEETHAADVOVIVIGLAARAFMOLERITLAAWFIIKOILYHTLAWFKEYPIME

CPn_0824 932677 932378
yscS/fliQ-YopS/fliQ Translocation Protein
IRTRAVLAFFATSFKSVLFEYSYQSLLLILIVSAPPIILASIVGIMVAIFQAATQIQEQT
FAFAVKLVVIFGTLMISGGWLSNMILRFAGQIFQNFYKWK

933618 932677 YSCR-YOP Translocation R
ERIKVFTIMRSIFRESLCFFTLSVSCCFADASLYENSCPSRCQPTPPPSNSNPLNVVQQP
VAASSVPSYMPPLNADDVLPROHLSDGSFSDTYPDITTQAIILIFLALSPFLVMLLTSYL
KIIITLVLLRNALGVQQTPPSQVLNGIALILSIYVMFPTGVAMYKDARKEIEANTIPQSL
FTAEGAETVFVALNKSKEPLRSFLIRNTPKAQIQSFYKISGKTFPSETRAHLTASDFVII
IPAFINGQIKNAFEIGVLIYLPFFVIDLVTANVLVAMQMMMLSPLSISLPLKLLLIVMVD

CPn_0826 934382 .933612
yscL-Yop Translocation L
HDNKRSGVFSSEVNOPORYYAIVKMKFFSLIFKDDDVSPNKKVLSPEAFSAFLDAKELLE
KTRADSEAYVAETEOKCAQIRQEAKDQGFKEGSESWSKQIAFLEEETKNLRIRVREALVP
LAIASVRKIIGKELELHPETIVSIISQALKELTQNKHIIISVNPKDLPLVEKSRPELKNI VEYADSLILTAKPDVTPGGCIIETEAGIINAQLDVQLDALEKAFSTILKAKNPVDEPSET SSSTDSSSLSNDQDKKE

935273 934434 CT560 hypothetical protein GCLVTANTFGTLDILMKHSKEDDLSRFLPKNLLVESPHPEEIPLKSLSFTMSWLPTIHPS WITIAMKEFPPEIQQQLLAWLPEPLVQEILPLLPGISIAPHRCAPFGAFYLLDMLSKKIR PCGITEEIFLPASSANAILYYTGPVKIALINCLGLYSIAKELKHILDKVVIERVKNALSP TEKLFLTYCQSHPMKHLETTNFLSSWTTDAELRQFVHKQGLEFLGKALTKENASFLWYFL RRLDVGRAYIVEQTLKTWYDHPYVDYFKSRLEQCMKVLVK

CPn_0828 936292 935267 yscJ-Yop Translocation J IKEYAWIMVRRSISFCLFFLMTLLCCTSCNSRSLIVHGLPGREANEIVVLLVSKGVAAOK LPQAAAATAGAATEGMADIAVESAU TTEALA ILNOAGLPRIKETSLLDIFAKGVAAQK
LPQAAAATAGAATEGMADIAVESAU TTEALA ILNOAGLPRIKETSLLDIFAKGULVSELU
CEKERYOEGLSEGMASTIRIMGGVVDASVOJISFTTENEBNILPLTASVYJKHRGVLDNPIS
IMVSKIKRLIASAVEGLVPENVSVVSDRAAYSDITINGPWGLTEEIDVVSVWGIILAKSS
LTTÄPILIFYVLILILEVISCGLLWVIWKTHTLIMTMGGTKGFFNPTPYTKNALEAKKAIG
AAABKEKKEDADSQGESKNAETSDKDSSDKDAPEGSNEIEGA

CP0€29 936729 937298 OS. CODUST homolog present in Genebank/EMBL as of 11/7/98
KYICFVPTLAKSFYINIRDSRFYSWLCFIMKETYYRDFLHENYLKNKKSMFWKIYKTAGE
FFLÄNAKWELVPAGYRRVRGKDFVLSPLVDLVILFFWVTKDSRYSFCSMTFTCICRS/VE
CIEWYSTLFGIGRFCAVWCVEGFSGSTFDKIYHTIVAVLGILGLGILTFILRIIFSYLML
PVWFLFKCYS

EDP_0830 937339 937959
NoFTOBUST homolog present in Genebank/EMBL as of 11/7/98
DSCSSLLPCFEVEAQTFPOVFSKVVVYKYKSSRILLIALLYNITLVIALLIFYHKKYLGOK
GRVILKIYONEEFFFAATERFPSIGAGYLRVENKSVLFPFEDLMLVCPSYFKOFPLSAF
KV*TIKLIYWSVLESIPVVGAFFFSIGRLFAMWCIEDFPGSIFSRIYHTIYGVLGILGLGI
IMFILRIIFTLLTLPFWLISCLKSSAA

CPT_0831 938249 938434
No LEOBUST homolog present in Genebank/EMBL as of 11/7/98
NKREANVLIRKSESEGAFFEATONYPTIQOGYOLVRIREHNLSVRAHFDLSLSLDASVHP

939750 .938827 CPI_0832

939750

938877

1ipA-Lipoate Synthetase

VMKCRPTLNTDQPRVRKKLPERFPKWLQRPLPQGSAFHATDATIKRSGMPTVCEEALCPN
RAECWSRKTATYLALGDVCTRSCGFCNIGHSKTPPALDPTEPERIALSAKELGLKHVVIT
MVARDDLEDGGAGGLVDIIOKLREELPOATTEVLASPFCNVSALHTLLDSGITIYNHNV
ETVARLSPLVRHKATYARSMFMLEQAANYLPDLKIKSGIMVGLGEMEGEVKOTLODLASI
GVRIVTIGGYLRPSRKHLQVKSYVTPETFDYYRRVGEMGLFVYAGPFVRSSFNADMILA SVODKASA

CPn_0833 :941171 939747

lpdA-Lipoamide Dehydrogenase
RGVLFEILITVSENMTQEFDCVVIGAGPSGYVAITAAQSKLRTALIEEDQAGGTCLNRG
CIPSKALIAGANVVSHIKHAEQFGIHVDGYT7DYPAMARKKNTVVQGIRGGLEGLIRSNK
ITVLKGTGSLVSSTEVKVIGODTT II KANH YILATGSEPRFPFGVPFSSRILSSTGILEL
EVLPKKLAIIGGGVIGCEFASLFHTLGVEYTVIEALDHILAVNNKEVSQTVTNKFTKQGI
RILITKASISAIEESONQVRITVNDOVEEBOYVLVAIGROFNTASIGLDNAGVIRDDRGVI
PVDETMRTNVPNIYAIGGITTGKWLAHVASHOGVIAAKNISGHHEVMDYSAIPSVIETHP
EIAMVGLSLQEAEQQNLPAKLTKFPFKAIGKAVALGAGDGFAAIVSHEITOOILGAYVIG
PHASSLIGEMTLAIRNELTLPCIYETYHAHPTLSEVWAEGALLATNHPLHFPPKS 941171 939747

CFn_0834 041544 CF556 hypothetical prote 942014 'n KTTOP TYPESTELLE TO FEET TO SEE THE PROBLEM OF THE TOP TO SEE THE PROBLEM OF THE

45698 942045 CPU_08.55 945098 942045

MOLL SWL75NF Limity helicidse
HNIMVLEALALFRODAMOHILEHRIKETVVDF/CEDSYTIKTEDERAPEGYWLSTLKLODID
RUFFASCSCHMISTCLILLMFANFAVYDALGLIFELIEKFFHCFW/AVFSHFFLDSTPLOAO
GEMYTLESPHITLIT BY LISBEVFODMERT HIASEE FYNTPINKFFLKSAL/RTAKKFFFL

CEMYTLESPHITLIT BY LISBEVFODMERT HIASEE FYNTPINKFFLER HEADTERWIGHT NEW ARTITOENE SEPTEMBELOW OF LVEKAR ILDEPTED FERLELAHTS LEWISHDI

DRENHPKTRIGSVEYVAKTHEMITGPKAIALPIYA
IPLLADKFKDOLLSLLCYDS RYDIRLLADASFSFSAYLVTPTDLDMGSLIYPMYC
YSPTKGLMQVVGMLSPKQAFIVKSEQVEDFLJÆRGHLIQEFGFOTF INERPEGKLTYNYT
EQGYLLFHYDVGDPSSTEIRFGTMTYYTNOGFFLEKKNDLPIQDGLIVEPDDIPAFIVKN
DAALRRLPNFFSSPPNLKDLLIEVHRQSROKGLDLKPILVGLGESRCWLFGVFLYREDI
FSLIPTPLQGLCFLPRVIPPENVPQFLTQYAQHERILFPNPQTRPPESYELVIQSIHRPH
PASPLHLQLELKTNLGSVPIGIALQGLJSKHTFLFTQAFLDLKQNLFQFLKGFLSTQKC
VIAENTVIANITDVFKLDALAPLSTYDDTIANPEDLOFFSQLKAACLPPIPQNLFSSDKC
VIAENTVIANITDVFKLDALAPLSTYDDTIANPEDLOFFSQLKAACLPPIPQNLFSSDKC
VIAENTVIANITDVFKLDALAPLSTYDDTIANPEDLOFFSQLKAACLPPIPQNLFSSDKC
VIAENTVIANITDVFKLDALAPLSTYDDTIANPEDLOFFSQLKAACLPPIPQNLFSSDKC
VIAENTVIANITDVFKLDALAPLSTYDDTIANPEDLOFFSQLKAACLPPIRQNSSKAFLTVTF
TOTAL TERMHILLSHILLATASTATETTPRINGLEDISTYTTLENNINGENSELTENNINGENSTALLSTYNGENSTALL THE THEORY THEORY THEORY TO THE TOTAL THEORY
CPn_0836 946960 945722
brng-Amino Acid Branched) Transport
KMKKNASHKTHDKKKSLSIWSIGGSIFAMFFGAGNIVFPLALGYHYNAHPWSAYFGMMLTA
VCVPLLGUSMLFYSGDYOKFFFSIGRIPGMIFITAIILLIGPPGGIPRAIAVSHATLIS
LSEHKSAFIPSLPJFSAICCVLIYIFSCKLSRLIGWLGSVFFPIMLVTLLWVIIRSFMIP
THPWVQEFIPNAGOAWLAGFIEGFNIMDLLAAFFFCSIVLISLRQLVAEEKHFTEEEIPJ
STGGISKKNRKSLALGFILAATLIGLIAATHAGLLUTVNSKGHILGRISAIALG
PNSILAGVSVPIACLITTEIALVGIVADFLARVVSFKKLNYASAVICTLIPTYLISILNFE
TISHLLLPLIQLSYPALIVLACGNIAYKLWNFRYSPVLFYLTLSLTIVLKLVN

CPn_0837 947145 nth-Enginuclease III LTMKQF/LRTLNALFPNPKPSLEGWSSPFQLLIAILLSGNSTDKAVNSVTPQLFAKAPDA OSILDLPPGKLYQLIAPCGLOERKSAYIYQLSQILVRDFHGEPPNDMALLTOLPGVGRKT ASVFLGIAYGKPTFPVDTHILRLAQRWKISEKKSPSAAEKDLARFFGHENTPKLHLQLIY YARQYCPALHHKIDNCPICSYLAKEANSTRT

CPn_0838 949196 947781 yndF-Thiophene/Furan Oxidation Protein ISLNIYPNSFHLFNLKLGILSESSFNFSIFMLKHDTIAAIATPPGEGSIAVVRLSGPOAI /ISLNIYPNSFHLFNLKLGILSESSFNFSIFMLKHDTIAAIATPPGEGSIAVVRLSGPQAI
VIADRIFSGSVASFASHTIHLGQVIFEETLIDQALLLLMRSPRSFTGEDVVEFQCHOGFF
ACSGILDALIALGARPALPGEFSQRAFLNGKIDLVQAEAIQNLIVAENIDAFRIAQTHFQ
GNFSKKIQEIHTLIIEALAFLEVLADFPEEGPDLLVPQEKIQNALHIVEDFISSFDEGQ
RLAQGTSLILAGKPNVGKSSLLNALLQKNRAIVTHIPGTTRDILEEGWLLQGKRIRLLDT
AGQRITDNDIEKEGIERALSAMEEADGILWVIDATQHEDLPKILFTKPSFLLWNKALD
PPPFLDTSLPQFAISAKTGEGLTQVKQALIQWMQKQEAGKTSKVFLVSSRHHMILQEVAR
CLKEAQQNLYLQPPEIIALELREALHSIGMLSGKEVTESILGEIFSKFCIGK

949230 950159 psdD-Phosphatidylserine Decarboxylase
FLFIVSRGLVQKPQYIDRITKKKVIEPIFYEKTMLFLYNSKLGKKLSVFLSTHPIFSRIY
GWLQRCSWTRRQIRPFWNRYKISEKELTKPVADFTSFNDFFTRKLKPEARPIVGGKEVFI
FPVDGRYLVYPNVSEFDKFIVKSKAFSLPKLLGDHELTKLYAMGSIVFARLAPFDYHRFI
FPCDCLPQKTRCVNGALFSVHPLAVKDNFILFCENKRTVTVLETEQFGNVLYLEVGAMNV GSIVQTFSPNQTYAKGDEKGFFAFGGSTVILLFLPNAIRFDNDLLKNSRMGFETRCLMGQ

CPI_0840 950141 951544
CT700 hypothetical protein
ISERRNLKTLKTFFGIAKRDKSQKWRIMWLVILWALAASLAIALVAKGYYRFVYFRRYAV
OVIREVRLSMELKEWALAEQOLDPILKKRSYRRQCLFEYMRILRKWGRFEESEKLLAEAK
KLGLRGPYFFLEIAYKAYRFGAFKECAQAFASVPQDLFEEDAAKYASALVRLGDLDAAC
SLIEPWISPLSHQETFVTMGHIYFTSKRYKDAIDFYNRANALGVCPVEVTYNLAQAYRIT
SSYAKAGKLFRKLLSNPVYKEEALFNIGLCEOKLGRPGKALLIYQSSDLWSKGDALLMKY
AAMAAMDORDVVLAEPCVELALRCSTFAKDYKCGLGYGFSLCRLRKYGDAERVYCNLIQN
FPECLTACKALAWLCGVGYATLLGSEGLMYAKKAVELDHSCETLELLSACEARCGNFDA
AYEIQSFLSSRDTSLQEKQRRSQILRILRKKLPLNDHHIVEVDALLAA 950141 951544

CPn_0841 951719 954640
seca-translocase Seca
IKRHMLGFLKRFFGSSQERILKKFQKLVDKVNIYDEMLTPLSDDELRNKTAELKQRYQNG
ESLDSMLPEAYGVVIONCRRLAGTEVEVSGYHQRWDMVPYDVQILGAIAMHKGFITEMQT
GEGKTLTAVMPLYLNALTGKPVHLVTVNDYLAQRDCEWGSVLRWLGLTTGVLVSGTLLE
KRKKTYCOLDVVGTASEFGFDVLRONSIATRLEEQWGRGGYFAIIDEVDSILIDEARTPL
IISGFOEKHNPVYFELKEKVASLVYLQKELCSRIALEARRGLDSFLDVDILPKDKKVLEG
ISEFCRSLWLVSKGMPLNRVLRRVREHPDLRAMIDKWDVYYHAEQMKEESLERLSILYII
VDEHNNDFELTDKGMQOWEYAGGSTEEFFVMMDGHEYALIENDETLSPADKINKKIAIS
EEDTLPKARAHGLRQLLRAQLLMERDVDYIVRDDQIVIIDEHTGRPQPGRRFSEGLHQAI
EAKEHVTIRKESGTLATVTLQNFFRELYEKLAGMTGTAITESREFKEIYNLYVLQVPTKKP
CLRIDHNDEFYMTEREKYHAIVNEIATIHGKGNPILVCTESVEVSEKLSRILRQNRIEHT
VLNAKJHAQEAEIIAGAKKLGAVTVATTMAGRGTDIKLDNEAVIVGGLHVIGTTRRQSRR
DPQLPGRCARLGDPGAKFFLSFEDRLMRLFASPKLITLIEHFRPPEGEMSDPMFNRL
IETAQKRVEGRNYTIRKHTLEYDDVMNKQRQAIWAFRHDVLHAESVFDLAKEILCHVSLM
VASLVMSDRQFKGWTLPNLEEWITSSFPIALNIEELRQLKDTDSIAEKIAAELIGPFQVR
FDHMYEZLSKAGGEELDASAICRDVKRSVMVMHIDEQWRIHLVDMDLLRSEVGURTVQOK
DPLLEFKHESFLLFESLIRDIRITIARHLFRLELTVEPNPRVNNVIPTVATSFHNNVNYG
PLELTT/TDSEDQD 951719 954640

CPn_0842 955015 954710 CT702 hypothetical protein (frame-shift with 0843) KYYTPPTISRSPWSNIALKTISEPEYDCNOLLKTOSILTTHVDTLLNAPKDFPNSKNOKH ILFCIAINTLSHYAQFLIAGNRRKFWIRYYNDOVMSEWTPFI

CPn_984: **843** 955330 954994 CT702 typothefical protein (frame-shift with 0843) NKNRLITOWREKYMYTDYDTPTYTHIDJTFEDLCDNSERASCLDKYOETTYYVEENTDQC DEL CAPAGEOALKMA LHORO

CEN_0844 847 956751 CEUTORAGE AT A GORDAL TOURS INDICATED IN ALCOHOLOGIA LE CANTON CONTROL AND CON OLEELSVORHEESEAALPSMTFPDFSEVFTEGFSPEER NVGKSSTINGLENEERCIIDNINGTTRDMIDILYSHKD WISSSRTEKAISRADICLLVIDATQKLSSYEKRILSLIS POOAPKTLKIALIGRP DTAGLRKMKSVKNSIE PHIILINKWDLLEEVRM

958151 956850 CPn_0845

CPD_0845

CPD_049 A E-Dymeto - DABSO

CARRO-FOR WRIALPECTER THE ALEXANI FLERING YOAKF MECUPDILIM

RPLEDIDIATNASPTIVSTIFPOUTSIGTAFGIT/WKQDGRLFEVATFRSDGYKDGRHP

DRIIFSSMREDALRRDFTVNGMYYDFEDKVFDFVEGTRDIEKKVIRAIGHPRLRFSEDK

LRILRAIRFSSSLGFTLDPTTERAIIKEAPALVNSVSPERIWQELKHUKRQPYGALSLL

LKLKVLIFIFPELRDIPYSLLRTTIEFARFNPTHFPEILFLLPLFQGVSEEAATVAFGR

RISNRELKIESWYEALPHFONGSONRVFWAHFLASPTAPLFLELFSALGKDPSRQOHF

ISRVQELESRLEGFILRIKTSSPVVSAPDLIAKGISPGRLLGDLLREAEILSIENECLDK EKILLLLOEKGFWK

CPn_0846 959383 958112

ClpX-CLP Protease ATPase
REHMKKNLTICSFCCRSEKDVEKLIAGPSVYICDYCIKLCSGILDKKPSSTISSAPVSE
TPSOPSOLRVLTPKEIKKHIDEYVIGQERAKKTIAVAVYNHYKRIRALLHNKQVSYGKSN
VLLLGPTGSGKTLIAKTLAKILDVPFTIADATTLTEAGYVGEDVENIVLRLLQAADYDVA
RAERGIIYIDEIDKIGRTTANVSITRDVSGEVQOALLKIVEGTTANVPPKGGRKPHOQE
YIRWITENILFIVGGFSVALDKIIAKRLGKTTIGFSDDQADLSQKTROHLLAKVETEDLI
AFGMIPEFVGRFNCIVNCEELSLDELVALLTEPTHAIVKQYMELFAEENVKLVFKKEALY AIAKKAKQAKTGARALGMILENLLRDLMFEIPSDPTVEAIHIQEDTIAENKAPIIIRRTP

960019 959387

CPT_U84/ 960019 959387

ClpP-CLP Protease Subunit

KLFDEETQMTLVPYVVEDTGRGERAMDIYSRLLKDRIVMIGQEITEPLANTVIAQLLFLM
SEDPKKDIQIFINSPGGYITAGLAIYDTIRFLGCDVNTYCIGQAASMGALLLSAGTKGKR
HALPHSRMMIHQPSGGIIGTSADIQLQAAEILTLKKHLANILSECTGQPVEKIIEDSERD
FFMGAEEAISYGLIDKVVTSAKETNKDTSST

CPn_0848 961556 960177

tig/murl-Trigger Factor-peptidyl-prolyl isomerase
v0assparpfksnkkgclvprslsnegfsvdleesfgctvsalvkvspevlnklnkQalk
kikkeitlpgerkgkapddvlasryptnvrkelgelvtodavhalstvgdrrelspkavk
sneiftofdloggakvefsyeafpaisdlpwenlslpogeaaseisdsdiekgliniom
atk#pverpsoggdpisislhvsksndenassaaifenkyfklseemfdafkekflgis
tgh#avvetitspetgsflrgdtutttvkavlevstpetddekarqlqaeslddlkaklri
QleMoakoklokkfseaedalahuvdfelpfsleeiistreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklinarlidycsdebelleeristitreklidycsdebell 961556 960177

THE THE TOTAL OF THE PROPERTY LPYGUEYALIANYHLLNSFYTDEYVDEXALLUSSTUDERDGAFRYKSEKKIVEFMTETIP LOYQOVRAFISDEGILARNILVEERKMLEEVFSGFTYDERDGAFRYKSEKKIVEFMTETIP ANGHRITFNCPENLSGOFIYDETIFELSFREGSDINYYEADLKVHGLLKGVPLDLLWDCI SAKKRFLELDKAGOQSKGTRRGKVNSGKLPCILVUDLEKIAPVVQIFNEIGFKVLDUSCI KCHEMSLTVISLDOFEALPVNFSMSERLIEIGKQIRGIEFDFODVPQQIGATLRSYQTE GVHULERLRKMHLNGILADDMGLGKTLQAIIAVTQSKLEKGSGCSLIVCPTSLVVNWKEE FRKFNPEFRTLVIDGVPSQRRKQLTALADRDVAITSYNLLQKDVELYKSFRFDYVVLDEA HHIKNRTTRNAKSVKMIQSDHRLILTGTPIENSLEELWSLFDFLMPGLLSSYDRFVGKYI RTCHYMGNKADNMVALKKKVSPFILRRMKEDVLKDLPPVSEILYHCHLTESQKELYQSYA RTUREMINKAUNVALKKAVSFT LIKRAKEUVENLEPVSELLINGHELESVELLISTA ASAMSELSRLVKQOEGERI HIHVLATITIEKKQICCHPAI FAKADAEPGDSAKYDMLMDLU SSIÄMESGHKTVVFSQYTKMLGI I KKROLESRG I PFVYLDGSTKNRLDLVNQFNEDPSLLVF LI SLKAGGTGLNLVGADTV I HYDMWNPAVENQATDRVHR I GQSRSVSSYKLVTLNTI E KILTLQNRKKSLVKKVINSDDEVVSKLTWEEVLELLQI

CPN_0850 965254 966390
mreB-Rod Shape Protein-Sugar Kinase
LGKKYMNCCRYDFMSPHRILFKLKNFSNRLYNRALGRFDKVFNFFSGNVGIDLGTANTLV
YVRGGGIVLSEPSVVAVDAQTHAVLAVGHKAKAMIGKTFRKIMAVRPMKDGVIADFEIAE
GMLKALIKRVTPSRSVFRPRILIAVPSGITGVEKRAVEDSALHAGAQEVILIEFPMAAAI
GVOLDYNEPAASMIDICAGTTEIAIISLGGIVESRSIRIAGDEFDECIINYMRTYNLM
IGPRTABEIKITIGSAYPLGDQELEMEVRGROQVAGLPITKRINSVETRECLAEPIQQII
ECVRLIEUK LOPEVERVENUV TGKALEHLDQFKKRKGNLV

CPn_0851 966378 968195 pckA-Phosphoenolpyruvate Carboxykinase REFGIVMWSTNIKHEGLKSWIDEVAKLTTPKDIRLCDGSDTEYDELQTLMESTGTMIRL REFGIVMWSTNIKHECLKSWIDEVAKLTTEKDIRLCDGSDTEYDELGTLMESTGTHIRL NPEHPHNCFLVRSSADDVARVEOFTFICTSTEAEAGPTNNWRDPOEMARELHQLFRGCMQ GRTLYIVPFCMGPLDSPFSIVOVELITDSPYVVCSMKIMTRWGDDVLKSLGTSGKFLKCLH SVGKPLSRGEADVSWPCNPKSMRIVHFQDDSSWMSFGSGYGGNALLGKKCVALRLASYMA KOGOWLAEHMLIGIG THE GEODIAWI REGGMEHMLIGIG THE GEODIAWI REGGREADHWISTERTSPASP PSAGGKTHLAMLMPRLGGKIELG GEODIAWI REGREDGELYAVNPEYGFFGVAPGTGERTNPNALATCRSNSIFTNYALTADGDVWWEGLTE QPPEPLTDWIDSKFWKFGGSPAAHPNSRFTFAPLRGCPSLDPEWNFOGVPLDAIIFGGRSETIFIGLTGCHSSTTTAALVGGLKLEHDEP AFLDFGGYMAVYFFGWML SPEAENBCLKLPKIFGVNWFRKNNGGEFLWPSFGENLPVLEWIFGRTDGLEDIAETPIGY LIPHINGHNLWSLNLDLQTVQELFGVDAEGWLAEVENTSEYLLFFGSDCFQOITDELLRIK GERLEK

968274 970613

CTT_UMS2 968274 970513
CTT/T1 hypothetical protein
TKLETDYYLIMTVTLOPSYINFTENVTTALSOKTLOBATELSCSALFFOELODKAQG
LKHALSUVOELSAKALRYAQVOTS CSYLETEESSRYSJOAGT FORTMETPTDDEVKATLQ
NEMFETOKTEVEGUBKVEKSYLDSVTPTEGTOPSNEYSAT FLNYTTLLNILKERFAASST
LTUADYHALYALIGDEVKETEALKAADAFPEKKNIAFYSETIMTYNNMOVLSYEVTDYLN
VOTABLISHATTAAOEVQYTLKNEYSTERDILHINSYTDEOMTHYPADAEYNARDAGVIQSL

LNLSGNYRQLTENMLPHTE I LAQIRG FOR IVNGTI LASHTUL PTTMEDTLLGV
LYTYQCCATI FOMSYGTSTP IDA I NQEKSYMOARANG FOVTODOVFOOFATNIQS
GTSYRG I DLFKNNKVNELIP I FLIQAASFLEK FYNLMSRSMYQT I EDAANRG ITALDGLI
SGWSTO I ATFOTOKNSLOFSLLKYFOTHKANKESFVTTA FLOWYYSSLMLDKYLPTQONV
LASLG I QMTYSNKAAKYLIELI KEITTFOSAD LYYSLISTLKQMILOAVADP I GKAVGVU
NDEKTRAMAD I TROKKIKAA I DINLVE I KADAELSKSQ I RELVOTLTNFKSQ SDDLI RNL
SCLLOFLSGLTLKAVND PHATYEAFTAE FTEPFNNWKRQLATFESFVI QGQONG I TPGG
OOQLLOAMESSQODFSTFICNOQ LALVESSAMQO EWTLVSAALALLIN MYSKIARRIKS

470637 971503

CTT.12 hypothetical proteid
NIHPKIEKRNSLPLTAVAPVFEES/HPSVATTVDYVDATTLSRHLTVLKDVIKEARNLD
LGKAFLTSMKOGF INTOTELAI IQASLADQSSRESRKLEKIFHQHLGKAAPQAATATSG
VOPTADPVADKHELGSAFAYVLLUKYIFAQEEALYALGREINLSGYAQNLFSPLLDHIKS
FNSAPINYNLGSYISQTSGTANFAVOYEMILSRYNNEVSQCRLDIASTVKAKAALANNSA
SVKANVSLTDACKKOIEDI IASYTKSLDVIHTQLTDVMTNLASITFVPGLNKYDPSYRIV
GGDLSI IALQNDEKVLVDGKVØITTAVNEGGLLNFFTTVLTDVQNYGDLAQTQQLMLDLE
LKAMQQQWSLVSASLKLLNGWTTVISGFKN

9/2849 971806 CPH_0854 972849 971806

ompB-Outer Membrane Protein B

GPFDMNSKMLKHLRLATLFFSMFFGIVSSPAVYALGAGNPAAPVLPGVNPEQTGWCAFQL

CNSYDLFAALAGSLKFGF/GDYVFSESAHITNVPVITSVTTSGTGTFPTITSTTKNVDFD

LNNSSISSSCVFATIALQETSPAATPLLDIAFTARVGGLKQYYRLPLNAYRDFTSNPLNA

ESEVTDGLIEVQSDYGTWAGLSLQKVLHKDGYSFVGVSADYRHGSSPINYIIVYNKANPE

IYPDATDGNLSYKEW\$ASIGISTYLNDYVLPYASVSIGNTSRKAPSDSFTELEKQFTNFK

FKIRKITNFDRVNFGFGTTCCISNNFYYSVEGRWGYQRAINITSGLQF

974001 CPn_0855 CPH_0855 974001 972994
gpda-Glycero1-3-P Dehydrogenase
GLMKQHIGYLGYGIUGFCLASLLANKGYPVVAWSRNPDLIKQLQEERRHPLAPNVVISPN
LSFTTDMKEAIMAFMIVEGVTSAGIRPVABQLKQITDLSVPFVITSKGIEONTGLLLSE
IMLEVLGDSVFPUGYLSGPSIAKEVLNGSPCSVVVSAYDSQTLKQIHEAFSLPTFRVYP
NTDIKGAALGALGALKWIATACGIAEGLSFGNNAKAGLVTRGLHEWRKLAAIMDCKPETLN
GLAGLGDLCYTCFSESSRNLRFGHLLAQGLTFEQAKAKIGMVVEGAYTALSAYQVAKHHK
IDMPITTGYRVLYENLDLKEGIALLLQRNTKEEFL

CPn_0856 975410 973995

AgX-1 Acmolog-UDP-Glucose Pyrophosphorylase
GSRDRVVRLTVMTESVYSPSAMMYNSLADKLKAINQEHILDIWPSLSPKQQQRLFQQLTS
VDIDBFRKQQQLLSSPTAILKDFHPITSFASSGEDPERAHAGTTLLKEKKVACVVLAGGQ
GSRLRCDGPKGLFPVSPIKKKPLFQLVAEKVRAASKLAGQPLPLAFMTSPLNTRQTRSFF
ESNJYFHLDPNQVDFFCQPLWPLLTLSGDLFLEDMDTLALGPNGNGCIATLLYTSGYWEK
WKNAGIEWYSVIPINDPLALPFDVELCGFHAMSNNEVTIKAALRQTAIEDVGILVKSWEK
GKTSVIEYSEIPQNERFALNEDGKLKYCLANIGLYCLSMDFIRHAAYQQLPLYKVHKHAK
QVGHTSLNEKNAWKFEEFIFDLFCYSDHCQTLVYPRQECFAPLKNLEGNNSPDTVRQALS
DRERQLFHKVTGKKLSPNTTFELEADFYYPSTSTSLHWENKAFFEEPFFEAS

975392 975808

CT7]6 hypothetical protein LLLAROYIKTARGISRLMRDRLGSLSLILKVKIHKYLDTLHNQKRLALTVSRNIQATNKR iaólhleryehfisronikhydilleylktlosslykogseslrfleihhogloelinrr Kiiekiknnkyskogeigt

CPn_0858 977115 975757

flii-Flagellum-specific ATP Synthase
RNSETRNQRRTRPSTFCFDSMNHLMKEKLHIHNWQPYRACGLLSKVSGNLIEVDGLSACL
GELCKISSTKDPNLLAEVIGFHNHTTLLMSLSPLHSVALGTEVLPLRRPPSLHLSDHLLG
GVLDAFGNPIDKKEDLPKTHRKPLLSLPPSPMMRQPIDQIFFTGIKAIDAFLTLGKGQRI
GVFSEPGSGKSSLLSAIALGSKSTINVIALIGERGREVREYIEKHSNALKQORTIIIAAP
AHETRAPPKVIAGRAAMTIAEYFREGGHEVLFIMDSLSRWIAALGEVALARGETLSAHQYA ASVFHHYSEFTERAGNNDKGSITALYAILYYPKHPDIFTDYLKSLLDGHFFLTSGKALA SPPIDILSSLSRSAQALALPHHYAAAERLRSLLKVYNEALDIIHLGAYTPGQDEELDKAV KLLPSIKAFLAQPLSSYCYLDNTLKQLEALADS

977597 CPn 0859 977055

CPT_0039
CT718 hypothetical protein
VFLVTTPQSPGSLSQSHLPHPHDPWDTEPTSLPEDPNDKASQELHSLVHLFRKLSIHLLS
EVEKTVQQLKPDLLELALLICEKFLYKKLENPQELALLLSTALQRHTTLRSLTPIKVFLH PEDLKTLTDWISTHELPMIKHAEFFPDTSCRRSGFKIETPNGILRQEISEELDHLLSVLT

978639

CPn_0860 978639 977608
flif-flagellar M-Ring Protein
rtLVFFONLAKKLTALGISFLGCLLIGGVVSCAILFGRSSNPSLAPTOVKTEKTSGNWLK
LTOMGNPKLIESLTKKECLEKDLTSFHPIASAKVAIALSTEDDVMSPLHLSVILTLRKEE SLTPSLLFSITDYLCSSLFGLKREHISLSDNLGNLYIPESITVNSLFIHTLENYLGKIFP KEHFALAYHAKAEKPTLQLTLNENYIAHLTKEESEKIVAHTKHYLYQNYDDSYDIVIETL PFARLQNKKSPPAKVLIGSMILVISLMIVALASFYLARHAYERVSPEPRKIKRGINISKL LEIIQKESPEKIALILSYLDPKKAEALLNRLPEDLKHQVLKYKL

CPn_0861 979752 978925 nifU-NifU-related protein ASYPPTWKFLMTLPLEPMIFWSSLSAK/MKKFLTPHCAGTFSEEDAEAKEAHLVTGKQGH ASTPT WAT BUT TEEPER PASSESAN YARRE LIFFLOALTS SEDARAR BARBANDA YARRA REMONCYTEYMUZDKKINGY ILDAKEQYFCHEYLI FLABAVCINLVGGKSYSSBAYKHTLDDI DKSLRYHAHOPALPEDSI ISLYHEVI DALDTAVEQCLE I PLEDGSLPLUNSPHINLDFEDAN PYSOSOMEALTHEOKU, YALRAT IAEK I GPY I AMDOGEVTVESLENF I VTI AYSONCSGCP SSLGSTLINSI GQLLRAY I YFELQWYDESSLINGHP

980924 ytho-mits-related protein GROTIFRITOCKTRCIRMENTONRKAPPIFWLNMOVAIPPSERVKERYALHSDIFSLPPG

GROTTE RITECULA, CHRISTOPHINAPP LEMONIOVA LEPOSERVES ALHEIDITE DEPO SALKLAEKTEES TROLUCERCHITERE PUBLIFICATI LULALUSERIA REMINI LI DAN DOOLLINSLORHOLI TITY CWATANHEYER LVEEY LI ETLEPERILIPERIA REMINI LI DAN LIDELLSLOKRIR LILHILDI SOLLIRA PLITTELLIA DI LITETA DALCOMOSI GOLTE TRISL ERMYSOWPE PHISANIA PERAVAMYTA PERE INALI PLETTETTERINI KIKLI OBLOSIVLESI GLARIS VONRLENIUWAATED PAREILA FILLIONS LYDILOY ERPOPLAOVLONG GOPE LCHOALHESUTERSKULEFSKLARAMHUA LKHUTPULGSSS

CPn_0863 **Q%/SD**081520 580831 pgmA Phosphoglycerate Mutace 980831 EHMALL ILLBHGOSVWNEKNLF3GWVDIPLSQOGIEE AIQNLPIDCIFTSTLVR SLMTALLAMTNHHSKKIPYIVHEDPKAKEMSRIYSAE PLYGSSALMERNYGELQ GKNKKQTAEOF3EERVLWRRSYKTAPPQGESLYDTKON JEYFENTLPOLQNGKNVFV SAHGNSLRSLIMDLEKLSEEEVLSLELPTGKPVVYQWKNHKIEKHPEFFG

CPn_0864 981658 982374
yjbC-predicted pseudouridine synthase
yGVNYTKVRLNKFLASATVASRRKCDEIIFSGSVTVNGRVAEGPFVL/DPEDKVOVGGTS
YGVNYTKVRLNKFLASATVASRRKCDEIIFSGSVTVNGRVAEGPFVL/DPEDKVOVGGTS
YMEDRICVYMMERA DIVINGL SIERPFFF TREEVIDLEAHLEYRVETVGFLOKETSGLILDVTH
DEFENKEIHHESSGTFFFZLAF/DPPMAKDEUFFLESTEIGERIVFF/DVTFFEFFTVK
IVVSEGKKHEIRLFADAAGFFILELKRIRIGSLVLGGLRYGEYRELTDAELGTYMKLSD

CPn_0865 982412 982942
CT865 hypothetical protein
SPMGYVFYVIAGSIFLGISLGAYCOLYYSVKSVLFSWYLLTVYALEKRHALLALSQLVGE
EDAQSQKEIDFLSQCDKLSWRAFLKNSYEIIPTFKEMEDLLSERVQGFLESIETIAEHDR
AILCIENFWASKNLFDFEIAAYEEAVEKYLKLRQRAPLRLASKLFRFLDVPSIRFSS

CPn_0866 983494 982916
bira-Biotin Synthetase
nMKVIYYEIEEIPSTNTMAKSYMHLWDPYALTVISTKCQTAGTGKFGKSWKSSKGDLLNT
FCFFITDLHIDVSRLFRLGTEAVVALCKOLGITEAKIKWPNDVLVHGEKLCGVLPETLPV
EGLLGVVLGIGLNKNTTKQALKDVGQPATSLQEILGHPIDLETTRELLIHHLLGVLQENL
PDSLARKSNRGNI

CPn_0867 983405 984667
rodA-Rod Shape Protein
CIRIPOMHIGFCHCVRGGNFFYFVINNFHILEIYSLLNSNTIMRYHKYFRYVNSWVFLVV
LITLMLLSVVVISSMDPTAMLVTSSKGLLTNKSIMQLRHFALGWVVFFICAYFDYHLFKRW
AWVLYFFMICALVGLFFVPSVGNVHRWYRIFFIHMSVQPSEYGKLVIVIMLSYILESRKA
DITSKTTAFLACLVVALPFFLILKEPDLGTALVLCPVTLTIFYLSNVHSLLVKPCTVVAT
GIIGSLLIFSGIVSNGKVKPYALKVIKEYQVERLSPSNHHQRASLISIGLGGIRGRGWK
TGEFAGRGWLPYGYTDSVFSALGEEFGLLGLLFTCLICFCCRTVAVATDDFGKLL
AAGITVYLAMHVLINISMMCGLLPITGVPLILISYGGSSVISTMASLGVLQSIYSHRFAK

CPn_0868 986733 984670

Znta/cada-Metal Transport P-type ATPase
NFRNGLGVRDLHHFREYVLI INEI LITCRYVFSRLFFTSFSAEVVNTFFESGMSEDTSPL
LSKORKKLSHNLPLKSAYLSLGTYLIALLSFWLHAKNLSNLFVVPTFFLAGTPALIKSLD
NICOKVVNIDILMTSAAFGSIFIGGALEGALLLVLFAISEALGOWYSKAKSTLVSLKQL
APTTGALVLEDNILOKVAINKI EVGNILRIKSGEVVPLDGETLHGSSSINLMHLTGEKVP
KSELHOSSIVPAGAHMMESSFDLRVLRTGSDSTIAHI INLVIQAQNSKPRLQORLDKYSSV
YALSIFAIACGIALLVPLFTSIPLLGPQSAFYRALAFLIRASPCALIFAITASINA
CANHGYLLKGGVILDRLVSCNSVVMDKTGTLTTGELTCIGCDYFGSKNETFFFSVLALEQ
SSSHPIAEAIVSYLMEQKVSSLPADRYLTVPGEGVRGYPNEQEAFVGRVPTSLGKVPSEY
LEDTEQKIVQAKOHGEICSLAYVGNSFALFYFRDIPROPAKRIJOQLKDGYVSMLTGD
HKVSAENTAEILGISEVFFDLTPEDKLAKIRELATORQIMMVGDGINDAPALAQATVGTA
KGERGSATAIEAADIVLLHDSLSSLPWIIQKAKQTKKVVSQNLALALAIILLVSWPASLG

CPP-0869 987479 986658
CT7228 hypothetical protein
ECWHFFFPKTSENTSDCRQHQILRKIMTQDPHDHFKSRTPEDHIKHVRDKHRVCKGEPHT
TFKGFFYHLANNALSTGVFIFFIRTLFFLIPTNRALQVKSLISLGVGWTFYHGCLKARKA
WAYWELSHRSMLEEKWEIEENFEQEKIELRILFENQGFKDPLLQEWVEYVCSDSTLLLDT
MIREELYIRKEDLPHPLIQGGSRILGGLGGLAIFLPLVLCISYTLAGVFSALMVLVLSFL
KAKILKNDKISEMVWVLGIFITSASIISSLMKLL

CPN_0870 988881 987448
set#F.Seryl tRNA Synthetase-2
TTHPTOGFGGAVILPFSPISIARRIKKSCCSEKSSIYSHFCTLLLINETSMLDIKIIRK
TPEECETRLRKKDPKISLEPVLSLDKEVRQUKTDGETLQAQRELLSQDIHKAKTOGVDAY
NLÎQEVETLAADLEKIEQHLDKNAQLHELLSHLPNYPADDIPVSEDKAGNQVIKSVGÜL
PIË#FPFKHLELNQELDILDFQAAAKTTGSCWPAYKNRGVLLEWALLTYMLQKQAAJGF
QLWLPPLLVKKEILFGSGIPKFDGQYYRVEDGEQYLVLIPTAEVVLNGFRSODILDEKE
LPHYWAACTPCFRREAGAAGAQERGLVRVHQFHKVEMFAFTTPNQDDIAYEKMLSIVEEM
LTELKLPYRLSLLSTDGMSFTASKTIDAEVMLPGQKAFYEVSSISQCTDFGSRRSGTTYK
DSQGKLQFVHTLMGSGLATPRLLVAILENNQQADGSVVIPEVLRPYLGGLEILLFKDQ

CPn_0871 988766 989899
ribD-Riboflavin Deaminase
EYMEDFSEQOLFFMRRAIEIGEKGRITAPPNPWVGCVVVQENRIIGEGFHAYAGGPHAEE
LAIQNASMPISGSDVVVSLEPCSHFGSCPPCANLLIKHKVSRVFVALVDPDFKVAGQGIA
MLROAGIQVVVGIGESEAQASLQPYLYQRTHNFPWTILKSAASVVGOVADSQGKSQWITC
PEAPHDVGKLRAESOAILVGSSRTVLSDDPWLTARQPQCMLYYKQPLRVVLDSRGSVPPTS
KVFPKTSPTLYVTTERCPENYIKVLDSLDVPVLLTESTPSGVDLHKVVFYLAQKKILQVL
VEGGTTLHTSLLKERFVNSLVLYSGPMILGDQKRPLVGVLGNLLESAPLTLKSSQILGN
SLKYVWEISPQVFEPIRN

CPn_0872 989903 991216
ribaribb-GTP Cyclohydratase & DHBP Synthase
kEPIFRVACLASESVNARESMIETREEVGSANFVSLEPAIEDLEAGKFVIVVDEASREDE
GDLIIAGEKTVEKMTFLLOHTTGVVCAALSGERLLSLDLPPM/KDNRCRFKTPFTVSVD
AAHGVTTGVSAADRTKVVQLLADPKSKPEDFISPGHFFPLAS*PGGVLKRAGHTESTVDL
MELAGLOPCGVLAELVNEDYSMMRLPQILEFARKHNIAVIPYTSIIAHRMLSDRLVSKIS
SAPLPTIYGDFTTHVVESLLEGMQHLALVKGNVAGKSNVLYSVHSECVTGDILGSKRCDC
GEQLSSAMSYIAEKGTGVLVYLRGQEGRGIGLGHKVPAYALQDNGYDTVDANLAMGFFVD
SREYGIGAQILVDLKLTTIKLITHNPQKYFGLQGFGLSIFERVPLPVRISEDNEQYLRTK
GEPMGHMLDLPCCNNRVQ

CHT_0874 993164 993744
CHT/3 hypothetical protein
EICHLIK (URKORDRERAMER (ER KKYLVFFLALLWOCHO (GYAGINNOLDYNDOTKVK
TOCETWI FOKLROYPELLWLTESOGAPLLOOTF (DMAYGEKLFNKKYFALDTATGMIHL

HLLIQUSROSYMQLSQILF TERROFOTAHIKOLLFFLNSFKSFENTLRILETARL RHVCCSAKAVTTFKPYFTOS STYAKALHVLRFFPELGTSYARLSPEQQEVLLSLRRL GNYDSLINLTE/PSAQLLSAMRTRSLAILDLYLVGLDTCUDKNCSGEFYINFAPLLSML QOHATIEEAFSPYFTYRANRLGFEGTSRTDMTL/RLATLMALGFSEASTLAWSFKNLPSD EAENLVNSFYTVQEHIPLTFRGLPSLVAGLSVATHGSTVSPENRLRQLYSTMLSLLVKS LRSHREMLNKQLLPQCTVLDFSETTLSSGGLVVFAESIAVRIHLNGAVSINL

CPn_0875 993363 994022

LET MERVI LYPT FOR BUILDED TO THE PROPERTY THE STATE OF SOME PROPERTY TO THE STATE OF SOME PR

CPn_0876

daga-D-Alanine/Glycine germease
SIATOETMILYFIEQUNKLSTSFOVFPHILLIGGFLTWKLRGLQFHGLKLGFNLMLQNKLD
DSSSKANEVSSYEAVAGILAGNFOTGNIAGMAVALACGSGALVWWLAALLGAIVQYAG
SYLGSKYRKPEGNTGEFIGGPIACLAFGMRKKILAGFFALFTIMTAFCAGNCVQVSCIVL
CARGOTFOKLLVGILLALVVJPVLAGGNNRILRFSARVIPFIAGFYCISCGIILFQHASA
ILPAIKLICSSAFGIKAGLAGIGGYTLSQVISTGINRAVMATDCGSGMYSILQANTKSKN
PVVDGLVTLVPPVIVMVVGITMLVLIVSGAYSGAQGTLMWASAFKNSLGSLGSVIVIL
MAALFGYTTILTWFACAEKSLQYMIPGRRANLWLKAIYVLIIPLGQVIDMRMIWALSDTG
FSGMVILNCIALIALLKDVLSTNRDVALLKERECSVADPVRNLDA

CPn_0878 996660 995992
SET Domain protein
GCMSTVTTEPCSSIHISLNNDWRDSQPYSLDRASELLHFRFLPSLVFSNWKVEQQIETLC
HKSEKRRLJSPLAWME,GKLHKQDLLCPPAPPVSVCWINAHVGYGVFARDETAPWTYIGEY
TGILRHRQAIWHDENDYCFRYPMPLFTLRYFTIDSGKQGNVTRFINHSEQPNAEAIGVFS
EGLFHVIJRTVAPIYAGQEICYHYGPLYWKHRKKREEFIPEEE

CPn_0979 997463 996645
yycJ_metal dependent hydrolase
YRILWKVSMQGFFPLASGSKGNSAYLGTDSCKILIDLGVSKQVVTRELLSMNIDPEDIQA
IFV/HEHSDHISGIKSFVKAYNTPIVCNLETARALCHLLDSHPEFKIFSTGSSFCFQDLE
VOYFNVPHDAVDPVAFIFHYREEKLGFCTDLGWVTSWITHELYDCDYLLIESNHSPELVR
QSQRPDVYKKRVLSKLGHISNQECGQLLQKIITPKLKKLYLAHLSTECNTAELALSTVSE
SKASITSIAPEIALAGGITSPIYFSRLEVACPR

CPn_0880 999864 997444

ftsk-cell Division Protein Ftsk

PHIRERKESRHPRLPTLPLAAKASLYLFFACFSGLSLWSFHRDOPCTONWIGLLGWSFSS

ELLYFFGAAAFFIPLYFLWLASLYYFRATFRPLFFYKAAAFLSLPFCSAILLSMLSPVGTL

PALLDTRLPKFILGNNPPVSYVGGIPFYLFYEQOSFCLKHLIGSVGTALITGFVMLFSVL

LYLGGIALLKKKTFODGVKKAFCSFFOTCFKNLKKLINRRNYLPKPSVPFVSKNPFSCTK

SØPSPRRVSETIILDGSISPLOGEIPGSKKESFFLTPHPCKRFLTKFVEPOENKAKEGK

TIALSSTPTVVMESKGKERAALPKLKSLAVPENDLPQYHLLSKNREARPESLOAELERKA

LILKOTLTSFGIDADLGNICSOPTLAAFEVLPHSGVKVQKIKSLENDIALKLQASSIRII
APIPGKAAVGIEIPTPFPQAVNFRDLLEDYQKTNRKLOIPLLIGKKANGDIMADLATHI
APIPGKAAVGIEIPTPFPQAVNFRDLLEDYQKTNRKLOIPLLIGKKANGDIMADLATHI
SREVYNALJWLVKEMESRYEILRYLGLRNIQAFNSTRNKTIEASVDRIRETMPFWGI
IDELSDLLSSSQDIETPIIRLAGMARAVGIHLILATQRPSREVITGLIKANFPSNISFK
VSNKVNSQIIIDEPGAENLMGNGDMLVLLPSVFGTIRAQGAYICDEDINKVIQDLCSRFP
TQYVIPSFHAFDDSDSDNSGEKDPLFAQAKTLILQTGNASTTFLQRKKKIGYARAASLID

CLEERIIGPSEGAAPROILLIONPLEG

CPn_0881 1005646 1006209
No robust homolog present in Genebank/EMBL as of 11/7/98
NKKFAVHMPVPIDNSSRNLQEVPESLEDLEQHAEESPTHQSAESSSLQLSLASSAISSRV
EQLSSLVLGMENSDFSSLRDVPIFSAIYESSTHTEVPTPLVCVGYINGSQSGYYDTQRES
LHLSQLLGSRRVEVVYNQGNFMEASLLNLCPRRPRRDPSPISLALLELWEAFFLEHPPGS

CPn_0882 1006169 1007404
No robust homolog present in Genebank/EMBL as of 11/7/98
NTPQVALLIQYFFCNGAFYVREALRLTPHAGNIVLVGICPSLYPEHPRSFYYRVSGDIGS
RFDDRGFVNSCVETLPYSSGSFGIFWISFTDPTFNFAIWITHRRTAGINEVSREMTODTE
TSLIEMROLSEQCEANNTDSLEQEESLMGIVGHTVGGVSMTVTSSPNIFYRIQTLLGLPE
TLAEAEENPTFFNSTIDSLAEIMMNLVRISDAVSIFWIFPIVDTTYNGVLLAVCIGFFGI
NGICSTFLMLTNPRSRRDRWRNLRIMVLCYRSLGSGMNLFDLSNNVRMARRHYTSCTVL
YAMVTLFGWTVAIQDALQYGFPSVRDAFYRYCLRHRYCLTQRNEDSLQTTGTRFQVTRT
HLEDQQMVASILNLSVFGLFFGFVGLMTTFGGLEISPSCRWDAANNRTVGIF

CPn_0883 . 1008904 1007573

dmpP/nqr6-Phenolhydrolase/NADH ubiquinone oxidoreductase
LYELFIKSGIFIVMTWLSGLYFICIASLIFCAIGVILAGVILLSRKLFIKVHPCKLKIND
NEELTKTVESGOTLLVSLLSSGIFIPSFCGGKATCKQCKVRVVKNADEPLETDRAG
QLEEGWRLSCCKVQHOMSLEIEERYLINASSMEGTVISNDNVATFIKELVVAVDPNKFIP
FKPGGYLQITVPSYKTNSSDWKQTMAPEYYSDWEHFHLFDQVIDNSQLPADSANKAYSLA
SYPAELPTIKFNIRIATPPFINGKNDELFIPWGVCSSYVFSLKRGDKITVSGPYGESFMKD
DDRPLIFILGGAGSSFGRSHILDLLHINHSKREIDLWYGARSLKENIYQEEYENLERQFF
NFHYHLVLSEPLPEDIAAGWDKDDPTKTNFLFRAFNLGQLGRLDNPEDYLYYVGGPPLHN
SSILKLGDYGWERSSIILDDFGS

CPG_0884 TOPESSER TOPESSER TOPESSER TALEFFY FILWEPEOKRR KAMEKEKNDLAKGOKYTAMGI ICTYDDI FEITY ILHI ACCEVEVLKGA I SETLEFNDNKS

CPn_0885 TO TO SO 1009433

yqcA - cRNA - Methy Tit ameteriaACTI PMCTMQNCPHEGOCGACHER (SAFYDDELEK PERLIHOJI FARIVICIONI ARTTECHE
DIRGENKMERSE FEGYYYESEN SIGERICE PERKOLI POPTY (LITHICYMD LLKITIKEMWIKH
PELMAYEPPKNKONI ATTI VICTORIO JAHRWI LITTO TE EVRVNEACTOEWKET LLCOOL

NIASTYWEEKVAARGISTYYETKILYGAPSTOOKLILK ASFSLAPRSFFOPOITO
AAKITETAFFINPEGSETLIDLYGGGGTIGIMLSPYL VEITPDAVASAQENIKA
NNKEDLVEVYLEDAKAFCKRNENKAPDVIITIPPRGG, VLKYILRIGSPKIVYISC
NPKTOFOECADLISGGYRIKKMOPIDOFPYSTHLENIILLEREIDP

CPn_0886 1011298 1010909
hcta-Histone-Like Developmental Protein
RTLEMALKDTAKKMKDLLDSIOHDLAKAEKGNKAAAGRVRTDSIKLEKVAKLYRKESIKA
RK.TGLARGPE TPALAFYRKTAEKFAEKFAEKTAAKSTERAFEAGFFAGKFAGKFAAKSTERAF

CPM_0887 1011692 1014157
CHLTR possible phosphoprotein
MKKLYHFTLFLRPLIRLSLIFALSLTLISGNFPOOKSFGHCCADMHSALISGKNCEELFA
MKKLYHFTLFLRPLIRLSLIFALSLTLISGNFPOOKSFGHCCADMHSALISGKNCEELFA
DFIERVLADRETLTARDMGTVVVU/REYLLKCIRKGDCDYGVKLLOKLLALRLPKDARKD
LOILMHRLNPEOAPLRDVVDOLFTIGCHESLODHLLFELYTVTLHSGYENRKODMLLAKE
GODYKKALELAKELVAALEKGSCSPHPEIVOIEKTFLOKTLLALOIKVAGEAGESCDALL
TPYCLSEIAYTEAMDALVLRIARGEVSRTNEVDSVLLSHALOHLPFAREKAIPELEVLID
HGAVLESTLLYYAFFSLLELYHONKOFASLERLLEKGDAVFVPEPHYPFPEGFFLGAFFY
HGAVLESTLLYYAFFSLLELYHONKOFASLERLLEKGDAVFVPEPHYPFPEGFFLGAFFY
KAGKYESSAEKVFLOIIDPAVKLGATFARAYEYLGCIAYVONHYEKAEEFYLLRAYKSVGRE
ESGIGLFLAYAVOKKKTACEDMLYHPKFSFTYRHLLDSLGSLSYPHGENKGSSAIQRVHR
ESGIGLFLAYAVOKKKTACEDMLYHPKFSFTYRHLLDSLGSLSYPHGENKGSSAIQRVHR
AVPELSEIJYSRCIYDMIKYRNVTYTHPIIELAYNOVRNLEKRNLEEICRDAQOPEYDKAL
AFWGALQSGASVPRSLIESSDVDEARITIRCYEALYFHNPDAIAMLPQAFSEECNSWOTA
LRLWWTLVREKGAPNHAKYWDHLVLRPHGDSLYFFGYDLOEVLIGKEDALKHLSVFALF
PKSSLLSLVYYLOGYSESSALRKVGWFYKALEEFTEISWSGEHMKTWAYIYYMWKLDLAD
TYISLGNFSQAVHILEEVKEDMQVASHPKLHELKGEDCYLAMELRWYEGLAYAYFQLHET
AHLSNHLLEHVEKNLISPRSYRDYYGGSLQRTLGLCQRFLGV

CPn_0888 1015441 1014119
hemG-protoporphyrinogen Oxidase
AERRFCVKRAI IIGAGISGLAAGMUHKKFPQAEILVLDKEAYAGGFVRTESPQGFSFDL
GPKGFLTRGDGEYTLKLIHELGLQNSLIFSDRAAKNRFVYYRGKAHKISTWTLLRKGLLP
SLIKDFRAPCYTQDSSVQDFLKRHSSQNFTSYLIDPLITAIRAGHSSILSTHMAFPELAK
REASSGSLLRSYLKNRSPKKSKTDRYLASLSPSMGTLITTIQEKLPATWKFSTSVTHIDC
SPKEACVTTPSETFFADWIYTGPLQQLFVLLPNYGIENLSKRVLPMILSSISLGMHAN
FSLPKGYGMLFADELPLLG IVWNSQLFPQATFGKTVLSLLIEGKWRESEAHAFAIAALSE
YLNINOKPDAFALFSSQDGMPQHAVGFLERKERILPHLPGNLKIVGQNIAGPGLNRCIAS
AYHAICDLHTEETLAQPQSSL

CPn_0889 1016841 1015462
hemN-Coproporphyrinogen III Oxidase
ffumrnykklecihopaprytsyptalewepsdaapallaforirenpoplslyfhipf
coswclycocsvvlnrrediveayintlioekklvvetigfrppvsrihfcogtpsrlsr
elffellforhikkifdlshaeeiaievdprslendmekadffonvoffnrvslgvodtoadv
Oeavrrosheeslkayekfkelafosinidliyolpkotkesfsktiodilamyporla
LF5Fasvymikphokamkasdmpsmeekfaiysoshilltkagyoaigmdhfslphopli
LAF6FMRTLIRNFGOYSLPPEDLIGIGMTST5FIRGIYLQNAKTLEEYHNTVLRGTFATV
KSKILTEDDRIRKWAIHKLMCTFTINKEEFFNLFGYEFDTYFIESRORLISMETTGLIHN
SESSLKVTPLGELFVRVIATAFDHYFLNKVSKKECFSASI

CHP-0890 1017829 1016819
heme-Uroporphyrinogen Decarboxylase
sthenwdsmaffdluksotashppiwllrougrymppyqelkgsqslktffhnteaive
athersellhydaaifadilsildgfavtydfapgprigfspeqpfftfsdpotifstiLD##RTLKQKLPYPLIVFAASPFTLACYLIDGGASKDFSKTMSFLYYYPENFOLISTIEGTAIYLKTQMDAGAAAVQLFESSELRESALFTRYVFENFRELIAKLKEQAIFVSLFGR
CFENFYTLQATQADTLHFDYHVDLHRIQKNLMLSLQGNLDPAIFLLPQEKLLHYVEAFL
V#LRTYPNFIFNSGHGILPETPLENVQLVVSYVQRQL

mfd-Transcription-Repair Coupling
NFMADFNPVNLDFSISKEFKETLPLLLENIHPGATAFLAAKMFHDCRASVIMITTPAR
LDDLFENLTFILDOAPVEFPSSEIDLSPKLVNIDAVGKRDHLLYSLNOHRAP LFCVTTLK
ALJEKTRSPQATSQOHLDLAVGDVLDPEATTELCKSLGYSOVMLTSEKGEFSGRGGIVDI
FPLSSFEFFRIEFMGEKIISIRSYNPSDQLSTGKVSKISISPAYTEASGGYNSHSLLDV
FSTEBLYLFDNLEILEDDFADISGTLSSLFDRFFSIGTLYDRISTSNQVYFSETPFPNVK
NLKBRRVIIEAFHRNMEASRQAIPTLYPEQIIQNDENPLLAFLOHLGETMPPHGKPLKLA
VSTENKSKLKEARALAETVARGDVEIYEKTGNLTSSFALVNEAFAAISLEFASTKVLRR
QKQRTHFSVTTEEVFVPIPGETVVHIHNGIGKFLGIEKKPNHLNIETDYLVLEYADKARL
VYPSNQAYLISRYVGTSDKAADLHHLNSSKWRSRDLTEKSLUYYAEYLLQLEAQRSTTP
AFVYPPHGESVIKFAETFPYEETPDQLKTIDDIYNDMMSPKLMDRLYGDAGFGKTEVIM
RAAVKAVCOGHROVIVMVPTTILATOHYETFKERMAGLPIEIAVLSAFSQAKVQKLICEQ
VASGQIDIIICTHKLINKSLEFKNPGLLIIDEGERFGVWKVKDNLKFRYPMIDCLTVSATP
IPRTLHMSLSGARDLSVIAMPPLDRLPVSTFVMEHNTETLTAALHELLRGQQAYVIHNR
IESIYTLAETIRNLIPEARIGVAHGQMGAEDLSNIFTKKNOKYTOILVATALIENGIDIP
NANTILIDHADKFGMADLYOMKGVGRWNKKAYCYFLVPHLDBLSGPAAKRLAALNKOEY
GGGMKIALHDLEIRGAGNILGTOOSGHIGTIGFNLYCKLLKKAVSALKKHTSPLLFNDDV
KIEFFYNSRIPDTYIETGSMRIEFYQKIGNAESSEELTAIQEEMRORFGPLPOEICULTE
GFLINAS

Opn_9893 1025888 rktB_TransketoTase EFLAFCLG LOYSCOFY LEG INKELLA TILLGK LAGALKQ IS LES LQKASSQHQL PLOCABLAAY LYGYVLRONF INKERFY JUGAGHGSALLYSCLHLAGFUVSLEDLQE PLOCABLAY LYGYVLRONF INKERFY JUGAGHGSALLYSCLHLAGFUVSLEDQUE FRQLHSKTPÖHPEYGETVOVEATTGPLOSTGNAVINAVLDGYLNEISVEDTKKFEAYG VCLAGDOCFMEGYSHEVCSFASSLALLNILWY TYDYNNVLDGYLNEISVEDTKKFEAYG WDVYEIDGYDFTH HETFSSI KROGERSTYLVIAHTI ICHGSPKEOTNRAHGSPLOVEDTH KKUPHLPEEKFFVPPAVKNFFAHK LØEDKRADEOWLDEVRVMSKOP PELHEEFVALTS EKKOPHLEEEKFFVPPAVKNFFAHK LØEDKRADEOWLDEVRVMSKOP PELHEEFVALTS HKLDKNLESLVOSVEMPDSI LAGRASANKLI LOVLVOH IPYLI LOGSADLSSDOTMIANSKV HTTYDFSGRNI KYGVBEFOMAT IMNILLYSOVFREFCOTFLVFSDYMRNATRLAALISKLE HTTYDFSGRNI TILMABE HETFOMAT INNILLYSOVFREFCOTFLVFSDYMRNATRLAALISKLE HTTYDFSGRNI TILMABHE HETFOMAT HAND LEGE BEEFTYTLFAT DELOTALLISKLE HELDKOVRVSPEPCMETER FALDVOVFYGOTTOGDEG REVSTEADSALDWYKY IGSEGLAIAMD RFGY SGASDDVSEECGFTTEQIJQRILSQ

CPn_0894 1036823 1025888
amn-AMP Nucleosidase
PRIDKNAKNILRRKHYKGEN/SKHTSESRIAQDMLERYSGSSVKQFCPYLLLTNFSYYIOT
FAKLHGVPVFEGSMFSAAJAPHLKTSILDFKLGSFGAALTIDLCSFLPDLKAALMLGMCG
GLBSHCVCDYFVPVAS/RGEGTSDAYFPPEVFALANFVVQKATTEVLEDKKANYHIGIT
HTTNIRFWEFNKKFRKKLYETKAQSAEMECATLFAAGYRRNLPIGALLLISDLPLRKEGI
KTKSSGNFIFNTYTEDHILTGQEVIENLEKVMLKRAASDHKKDQQYRGLPHMEVGEADDT
MAGGSETSDSDY

CPn_0895 1026973 1027557
efp-elongation Factor P
EIDCFMWRVST95FRVGLRIEIDGOPYLILONDFVKPGKGQAFNRIKVKNFLTGRVIERT
YKSGESVETADIVERSMRLLYTDQEGATFMDDETFEQEVVFWEKLENIROMLLEDTIVTL
VLYNGDVVAVPPIFMELSIAETAPGVRGDTASGRVLKPAVTNTGAKIMVPIFIDEGELV
KVDTRTGSYZSRVSK

CPn_0896 1027574 1027822 CTT53 hypothetical protein EKYPFFTVRNMEAKKIKELSKEAQLLKKLREKSRVLDEKNKRKAWVAKLVAMPESIREIE KEERVTTPQLFQAIAEKILEEGV

CPN_0897 1028794 1027853
(ghosphohydrolase)
nFLDSNTVDQKNKSNPRPMOEKPRHVHRIIHISDVHFHVLPVNPVHCFNKRLKGLLRKV
NFLVHFQATTIGQRFPKVVRSLGADSVCITGDFSLTAMDGEFLLAKHFVETLAKHSSVYL
LPGNHDVYTLKSLAQOTFYTHFPNDQLQQNKVSFHKITDHWALILLDCSCLNGWFSANGV
VHLAQISAIETFLLSLSPEENVIIAHYPLLSSQNPSHDLINNTHLQNVLKKYPKVRLYL
HGHENQAAVYNCADTSPSYILNSGSISLPTNSRFHVIDLYPEKYQVHTMILKNLLDFDAP
LEIANEATWDCQKL

CPn_0898 1030511 1028904
Mitochondrial HSP60 Chaperonin Homolog
TKKRIGSVKILRLLGVCHSEDEKLSNYMADKKLFSGIDKLFQIVKGSYGPKQSLSPTSFF
TKKRIGSVKILRLLGVCHSEDEKLSNYMADKKLFSGIDKLFQIVKGSYGPKQSLSPTSFF
KERGFYAISQTELSHSYENLGVDFAKAMVNKIHKEHSDGATTGLILLHAILQESYAALEK
GISTHKLIASLKLQGEKLDEALQQQSWPIKDALKVRNIIFSSLHMETIADHFYNAFSVVG
PEGLISITKERENDKTSMDVFQGFKIPAGYASTYFVSDTASRLTRIAHPLILITDRRISH
HISLLPLLQEISEQNQHLIIFCEDIDPDVLATLVVNKLQGLLQVTVVTIPQLSTTNQELA
EDIALFTCTHICPCQEASHVLAPEMVTLGSCLSIEISESQTTLIGGLHIPEVLTLKTRQL
AEEIRTTSCLETKKRLIKSTRRLQSSVAILPTDEDNEPLYTLALKIMESALSRGYVPGGG
VALFYASLTLGTPKDDADENSIAISLLQKACCAPLKLLATNADLDGDAVIAKLSSLGTTS
LGISVFSREIEDLIAGGILDSLATTSTILAQALDTAILVLSSKILILENQYEISTL

CPn_0899 1030848 1032215
muff-Muramoy1-DAP Ligase
MHRCCRQMYMRAMLLEDAVSIMLSDVSCPKCDKKITGFAIDSQQVQPGDLFFALPGNATD
GHQFLKHAATAGAVAAVVSHDYQGDSFGLELIRVDDTKSALQEAGSNQCNLFCOTLVGIT
GSVCKTTTKEFSKTILSSIYKTHASPKSYNSQLTVPLSLLMAEGDEDVMILEMCVSEPGN
MQDLLRIVQPEIAVITHINDQHAMHFPOGIGEILKEKSYILQKSKLQLLPKDSPYYLDLR
SCSPTAEKFSFSFNDPLADFCYKAISGDSVVIQTEEENYCLPIAFSYKPAYTNLLIAVAL
SWILEVPEEGVIRSLPELKLPPMRFEHSMRNGMQVINDAYNACPEAMIAALDALPLPSOK
GKIILIIGHMAELGRYSEEGHALVAEKAASRGDMIFFIGEKWIPVQSVLKSYSCEVSFFS
SAQDVKDILKQVARYGDVILLKGSRALALESLLACF

CPn_0900 1032208 1033281
mray-muramoyl-Pentapeptide Transferase
LVFNFLGASMIPLIPMFLKQSLFFSLALTGMTTLVLTVALGVPVMKWLKRKNYRDYIHKE
YCEKLEMLHKDKAEVPFGGGVLLF1SLIASLLVMLPWGKFSTWFFIILLTCYAGLGWYDD
RIKIKRKQGHGLKAKHKFMVQIAIAAFTLIALPYIYGSTEPLWTLKIFPMEGMLSLPFWL
GKVFCLGLALVAIIGTSNAVNLTDGLDGLAGTMSFAALGFIFVALRSSTIFIAQDVAYV
LAALVGAGCIGFLWYNGFPAQLFMGDTGSLLLGGLLGSCAVWLRAECILVVIGGVFVAEAG
SVILQVLSCRLRKKRLFLCSPLHHHYEYQGLPETKIVMRFWIFSFVCAGLGIAAVLWR

CPn_0901 1033239 1034537
murD-Muramoylalanine-Glutamate Ligase
FCMRRSRYSGCLMEIDMCQRILILGTGITGKSVARFLVQQGHYLIGADNSLESLISVDHL
HDRLLMGASEFPENIDLVIRSPGIKPYHPWVEQAVSLKIPVVTDIQVALKTPEFQRYPSF
GITGSNGKTTTTLFLTHLLINTLGIPAIAMGNIGLPILDHMQQPOVRVVEISSFQLATQE
HIPALSGSVFLNFSRNHLDYHRNLDAYFDAKLRIQKCLRQDKTHVWEEGSLGNSYQIYS
ELTEELLDKGDALKPIYLHDRDNYCAAYALANEVGWVSPEGFLKAIRTFEKPAHRLEYLG
KKDCVHYINDSKATTVTAVEKALMAVGKDVIVILGGKDKGGDFPALASVLSQTTKHVIAM
GECRQTIADALSEKIPLTLSKDLQEAVSIAQTIAQEDTVLLSPGCASFDQFQSFKERGA
YFKLLIREMQAVR

CPR_0902 1034507 1035241
nlpD-Muramidase (invasin repeat tamily)
AVDQRNASSEVNANDROMYITAVVANAILLUALEYTSKRIGVKDYDEGFPNFASSKVTQA
VVSEEKVIERPVASEVERBIAKETLAAGFIESKEVIVTTPPVPVVSETPEVPTVAVPPQ
PVRETYKEEQAPYATVVVKKGDFLERIAPANHTTVAKLMOINDLITTOLKIGQVIKVPTS
QDWSNEKTPCYCYTANFENYYIVQEYDGSWTIALRNHIRLDDLLKMNDLDEYKARRLKRGD
QLRIR

10355 Y

CPG_0903 TO 0000 TO 00

SSLIANMCGVTLLLKV

CPn_0904 1035320 1037396
mucG-Peptidoqlycan Transferase
RYMMKKIRKVALAVGGGGGHIVPALSVKEAFSREGIDVLLLGKGLKNHPSLQQGISYREI
RYMMKKIRKVALAVGGGGGHIVPALSVKEAFSREGIDVLLLGKGLKNHPSLQQGISYREI
RYMMKKIRKVALAVGGGGGHIVPALSVKEAFSREGIDVLLLGKGLKNHPSLQQGISYREI
RYMMKKIRKVALAVGGGRAVARTIGUNGGVARAVERANTIGUNGGSVALARVERANTIGUNGGSVALARVERANTIGUNGGGATIAKERS
LETTELOVA DES JAAN LETTE JARVERSI DEN DEN KONTALLER FOR TENTAL JOHN DE LANGEN DE STANFORMENT AN AVEG DE STANFORMENT AND AVEG DE STANFORMENT DVLEGGTMILEKELTEKLLVEKVTFALDSHNREKQRNSLAAYSQQRSTKTFHAFICECL

CPn_0905 1037400 1039835

murCsddla-Muramate-Ala Ligase & D-Ala-D-Alam Ligase
VHYMKGTPQYHFIGIOGIGMSALAHILLDRGYEVSGSDLYESYTIESLKAKGARCFSGHD
VHYMKGTPQYHFIGIOGIGMSALAHILLDRGYEVSGSDLYESYTIESLKAKGARCFSGHD
VHYMKGTPQYHFIGIOGIGMSALAHILLDRGYEVSGSDLYESYTIESLKAKGARCFSGHD
VHYMKGTPQYHFIGIOGIGAANCLNGYSGSSKIFVAEADESDOSLKHYTPRAVVIT
NIDNEHLNNYAGNLONLVQVIQDFSRKYTDLNKVFYNGDCPILKGNVQGISYGYSPEOL
HIVSYNQKAWQSHFSFTFLGQEYQDIEINLPGQHNAANAAACGVALTGIDINIIRKAL
KKFSGVHRRLERKNISESFLFLEDYAHHPVEVAHTLRSVRDAVGLRRVIAIFQPHRFSRL
EECLOTFPKAFQEADEVILTDVYSAGESPRESIILSDLAEQIRKSSYVHCCYYPHGDIV
YLRNYIRHDVCVSLGAGNIYYIGEALKOFNFKKLSIGLVCGGKSCEHDISLLSAQHVSK
YLSPEFYDVSYFIINRQGLWRTGKDFPHLIEETQGDSPLSSEIASALAKVDCLFPVLHGF
FGEDDTIOGFFEILGKPYAGPSLSLAATAMDKLLTKRIASAVGYVVYPYQPLNLCFWKRN
PELCIONLIETFSFFMIVKTAHLGSSIGIFLVRDKRELQEKISEAFLYDTDVFVEESRLG
SREIEVSCIGHSSSWYCMAGPNERCGASGFIDYOEKYGFDGIDCAKISFDLQLSQESLDC
VRELAERVYRANGGGSARIDFFLDEEGNYWLSEVNPIPGMTAASPFLQAFVHAGWTQEQ
IVDHFIIDALHKFDKQOTIEQAFTKEQDLVKR

1040514 1039915 CPH_0906
CTT63 hypothetical protein
kWGSEVLELVNDSQLSREASAFRLDIDFFILNIYPFFRNFKNIELCFFLSISQFNLDFME
EFVAYIVKNLVTNPEAVEIRSIEDEDNESIKLEIRVAAEDIGKIIGRRGNTIHALRTILR
RVCSRLKKKVQIDLVQPENSTDVIADQDYICDNDSSNSTEDTFGESDTCCSGHCHYDEDL NOEEOEEGNMHHSCECSNHH

CPn_0907 1040816 1040445

*CutA Periplasmic Divalent Cation Tolerance Protein CutA (C-Type Cytochrome Biogenesis Protein)

FAFSKFLIIKSSMTAVLILTSFPSEESARSLARHLITERLASCVHVFPKGTSTYLWEGKL
CESEEHHIQIKSIDIRFSEICLAIQEFSGYEVPEVLLFPIENGDPRYLNWLTILSYPEKP
DI SP:

CPD_1908 1041607 1040780
CT764 hypothetical protein
ILAT FMILIKUNELMIRRFFKTLFPPGPQYSLCYASILIVLSSLVCVPTFCWLFLPELS
LSKTMPSPIRNLFLVSSTLSKVPPTAIAEHLRLSADAPTYLHEFSIKEAESSLHALGIFS
SLYZEKSPDNKGITIFYTLOTPIAYVGNRSNTLCNLEGSCFLGQPYFPSLNLPQIFFSQE
DLGGGKLPKEKMLFTKILLKELAMESPKIIDLSLSDAYPGEIIVTLSSGSLLRLPIKTLD
RALDLYKHMKKSPVIESEKQYVYDLRFPNFLLLKAL

CPA-0909 1041592 1041966
rsby-Sigma Factor Regulator
IligistrrfllerlimnlsakeygdliviylogsldavsvpsvQeyleQfiQkkHlkial
NPTDYSYISSAGIRLLLSNFKLVQSLGGKMCLCCVKESVTEVMRIAGLDQLILLCQSEQE
CLSKL

CPD 0910 1041970 1043004
miāmitrna pyrophosphate Transferase
FLYMLPFEFENTTSSPECOVCLDPOKLFYKLFKRTIVLLSGPTGSGKTDVSLALAPMID
GEIVSVDSMOVYQGMDIGTAKVSLKARĢEIPHKLIDIRHVQEPFNVVDFYYRAIQACONIJ
LSRNKVPILVGGSGFYFHAFLSGPPKGPAADPQIREQLEAIAEEHGVSALYBULLKDPE
VÄÖTİTKNDKNKIIRGLEIIQLIGKKVSOHEMDIVPKASREYCCRAWFLSPETEFLKNDI
QMRCEAMLQEGLLEEVRGLLNQGIRENPSAFKAIGYREWIEFLDNGEKLEEYEETKRYFV
SNSMYTKKQKTWFKRYSIFRELPTLGLSSDAIAQKIAKDYLLYS

CPn 0911 1044079 1042985 CPH-0911 1044079 1042985

Fe.5 Cluster oxidoreductase
SLEAIFNANYFMALCKRISFEEGLELFVSSPIERLQERADAIRKERYPSNEVTYVLDAN
PNYTNICKIDCTFCAFYRKPKSPDAYLLSFDEVRSLLQRYVSSGYKTVLLQGGYHFGLGI
DYLEELVRITYQEFPSIHPHFFSAVEIEHACRVSGISTEGLQRLWDAGQRTLFQGGAEL
SERVRKIISPKKMQPGGWINLHKLAHLMGFRTTATMMFGHVENPEDILIHLØTLRDAQI
SCRGFYSFIEWSYKPONTALRRNVPQQASIETYYRILALGRIFLDNFDHVAASWFGEGKS
LGAKALHYGADDFGGVILDESVHKATGWSIQSSEEEICNIIRSEGFIEVERNTFYQHISC

1044120 1045760 CPH_0912 1044120 1045760
CTT68 hypothetical protein
VVIMDNSDNSFHTLETEQGSFLNDELAVEEVASTESTEISDATLCBAEKKVAFILNKMRE
ALMOSSQSOLRLFWOLRKOCPLFNREIEDTAKRADHWRCYIELTKEGRHLKGLQDEEGS
FVVQOIDLAITCLEKDILKFQEGTEDKIFKDREDNFLESQALDKAQAFYKQHHTSLLWLS
SFSSKIIDLRKELINVOMRNRLKSKFFORLSNLGNQVFPKRKEJIEKVSOFFAEDVDAFV
AKYPIGSBKETLKKTVFFLRKEIKNLGNAKRLFVSSHVFAEFTRIKLSKKUPOLKGMEKE
IRQEQGRLRVVSAENSKEVRQMLAEVSSLLIEGNDLSKVRKYLEGISKKIRALDLTHDDV
LGLKKENQOLFDOLREKQDAAEHSYQEQLAKDKQVKKEAASBLAERITTFSKTSEGNIT
SESREEWOTLKELLGKMSFLPPEKISLDNQLNLALQTIVMFFEEQLLSSPDSREKLVMM
RQVLKORRERQELKDKLEQDKKLLGSSGLDFDRAMQYSALVEEDKRALEELDASILELK OUTOOLL

CPn_0913 1045709 1045945
No robust homolog present in Genebank/EMBL as of 11/7798
RICKYPRIEATDSALAMRNCIYAFDLDGTLLKUNGSWSFYCYGLLQGLFSYKTLPPCI
YRFFREKFFFGIFHPSIIR

CTO_0914 104999 104639 NO CONSIST HOMOTOG DESIGNE IN GENERALIZEMBLI AS OF 1177798 VERWOLESETYS UTTICLES/VEDDLYEVALUEVATLIT/SIDEYARVLEKLEEAFADTTYO VELES/SISTOR FUND LAQUET SOWYASSYRD/SAEZET LYKKELTYDKKAQ ELSYTKKENQ ARGITTE/ODDLELOLDFEMILGEEKTVVRPQGYLKKMAXYYYWN EV

1046401 3046817 CPn 0915

YBBB LOJAP SUPERLAM LANGUAGE FOLKKASTUSKI LEYPKAGE FOLKVAAK TODKKONNUVULOVRTISEFTOYFV FVEGSVNVHVKALAIT IVEELKKOKVSFLIVEG JEDGMVVIDYGF IVVHVFVSE IRGKY RLEELWKDGFIVTSKLLAS

CPn_0916 1046813 1048984

£abF-Acyl Carrier Protein Synthase
LLNGVRVYMSKKRVVVTGEGV/SCLONEVCFFYDNLLAGVSGVRPITSFPCEDYATRFAG
LLNGVRVYMSKRVVVTGEGV/SCLONEVCFFYDNLLAGVSGVRPITSFPCEDYATRFAG
SELBELSE MERELS HERELSER FINALLY FRANCES HERELSER FOR TALATORECI DAYOHUVSORADMI ICCGTEAAVNRIC/LEGFIANRALI MENERI MILETVIL TAMATSINEST DAYOHUVSORADMI ICCGTEAAVNRIC/LEGFIANRALISERNDAPDQASREMDRADGV LGEGAG ILVLETLESALRRDAP IF ABH/LSYVTCDAFHITAPRDDGBGITACVLGALNSA GIPKERVNYVNAHGTSTPLGDLSEVLAVKKAFGSHVRNLRMISTKSLIGHCLGAAGGVEA VVAIQAILTGKLHPTINLDNPIAETEDFDVVANKAQDMDIDVAMSNSFGFOGHNSTILFS

1048064 1048539 CPn_0917 1048054 1048539
hydrolase/phosphatase nomolog
frddilevctlvmmkrkytysfydlyflkffotpdkntlkacfichtrgkhwgfpkghsed
keopoeaaerelveetglsvynffpkvliegysfnneegyfvrkevtyflaevrgdihad
pmeicdsgwlsloeglrllsfpelrdltveadkfinnylfss CPn_0917

049232 1048579 CPn 0918 CPH_0918

ppa-Inorganic Pyrophosphatase
ELLMSKKPLYVAHPWH9PTLTODNYESLCCYJEITPYDSVKFELDKATGLLKVDRPQKFS
ELLMSKKPLYVAHPWH9PTLTODNYESLCCYJEITPYDSVKFELDKATGLLKVDRPQKFS
NFCPCLYGLLPQTYCQTASGNYSGEQTRREGIQGKDPLDVCVLTEKNIHHGNILLQARP
IGGLRIIDSGEADDKJIAVLEDDLVFABIEDISDCPGTVLDMIQHYFLTYKATPNHLIKG
SPAKIEIVGIYGKKFAQKVIQLAHEDYLSYIGDTAEVN

1049375 1050430 CPn 0919 CPn_0919 1049375 1050430

1dh-Leucine Dehydrogenase
FMYSLNFKEIKIDDYERVIEVTCSKVRLHAIIAIHQTAVGPALGGVRASLYSSFEDACT
PMYSLNFKEIKIDDYERVIEVTCSKVRLHAIIAIHQTAVGPALGGVRASLYSSFEDACT
DALRLARGMTYKALISNTGTGGGKSVIILPQDAPSLTEDMLRAFGQAVNALEGTYICAED
LGVSINDISTVAEETPYVCGIADVSGDPSIYTAHGGFLCIKETAKYLMGSSSLRGKKIAI
GGIGSVGRBLLQSLFFEGAELVVADVLERAVQDARRLYGATIVPTEEIHALECDIFSPCA
RGNVIRKDILADLNCKAIVGVANNQLEDSSAGMMLHERGILYGPDYLVNAGGLLNVAAAI
EGRVYAPREVLLKVEELPIVLSKLYNQSKTTGKDLVALSDSFVEDKLLAYTS

CPN_0920 1051423 1050431

CysO_Sulfite Synthesis/biphosphate phosphatase

ILEENSHHSELPNYONIVESVVTEITTOLLNYRSEHRLVPFWEKSDGSFITAADYGSQYY

LKOLAKAFPNIPFIGEETLYPDODNEKIPEILKFTRLLTSSYRDDLISTLVPPPSPTS

LFYLVDPIDGTAGFIRHRAFAVAISLIYYEYPEILSWACPAYNOTFKLYSAAKGHGLSIV

HSONLDRRFVYADRKOTKOFCEASLAALNOOHHATRKLSLGLPNTPSPRRVESOYKYALV

AEGAVDFFIRYPFIDSPARAWDHVPCAFLVEEAGGRVTDALGAPLEYRKESLVLNNHAVI

LASGDOETHETTLAALONQLNVVPTDKLIAL

1051526 1052293 CPH_U921
snGlycerol-3-P_ACyltransferase
GELMLIKLMRATYEGMYTFLVGALLKLRYRMQVEGMDTLNINPKQGCLFLANHVAEVDPI
ILEYLFWSRFHVRPMAVEYLFHSRVVQMFLHSVRSIPIPQLVPGKESKRSLERMNVCYEE
ASRALNRGSELLLYPSGRLSRTGKEEIVNQYSAYVLLHRVMECNVVLVRVSGLMGSAFSR
YKONSTPKLGPAFKEAFRALLRRGIFFMPKRFVKITLCQVDHLFLKQFPTKQDLNTFLAS

CPn_0922 1052266 1053927

aas-Acylglycerophosphoethanolamine Acyltransferase
OFAHRSLRITRKLRRHHDORNRGHNNINLRLRPGSTLLEAFLILCSEHEEGIACFDEHL
GSLSYRELRNAIIAVAIKVSKFSEDRVGVMMPASIGAFIAYFGILLAGKTPVMMNMSGL
RELRACTKTVEVRRVLTSQOFIKHLTEVQGFVEYFFDIMYMEDVRKRLSWMEKCRIGLY
KCSVPMLLRIFGVSGVESDDTAVILFTSGTEKLPKAVPLTHKNLMENQEACLKFFDRNTO
DVMLAFLPPFHAYGFNSCGLFPLLMGVHVVFASNPLNPKKLVEFIDDKKVTFFGSTEVFF
DYILKTAKKONSCLESLRLVVIGGDALKDTLYEETKKLQPQIALYQGYGATECSPVIST
TKESPRKSECVGMPIEGEMDVLIISKETHIPVSSGEQGLIVVRGNSVFSGYLGNHEHQSFV
SLGGDQWYLTGDLGHIGPSGDLFLEGRLSRFVKIGGEMVSLEALESILHEHFTENQNEDA
GSLVVCGIPGDKVRLCLFTTLATTIHEVNDILKSAETSSIVKISYVHQVESIPILGIGKP
DYVSLNALAUSSLFG 1052266 1053927 DYVSLNALAVSLFG

1053966 1055093 CPn_0923 1053966 1055093
biof_1-Oxononanoate Synthase_1
vKESFLTTSDVIDFVTMDFLGFARSPTIYCEVSKRFQIHCQOFPHEKLGIRGSRLMVGP
SSVIDDLESKIASYHGAPNAFIVNSGYMANLGLCHHVSRSTDVLLMDEEVHMSVVHSLSA
ISGOHHTFHHNNLEHLESLLQCYRISSKGRIFIFVSSVYSFRGTLAPLEQIIALSKKYHA
HLIVDEAHAMGIFGDDGKGLCHALGYENFYAVLVTYGKALGTMGASLLTSSEVKYDLMQN
SPPLRYSTSLSPHTLISIGTAYDFLASEGEIARKQVFKLKEHFHECFDSHAPGCVQPIFL
PHTCLEEAISVLETTGIHVGVVAFAKHPFLRVNLHAYNTVDEVNLLAQVMKPYLEKSSHR VHINHEFHLWRELCCH

CPO_0924 1057301 1055028

DITA-Primosomal Protein N'
KRFTAKTKSMGYIESSTFRLYAEVIVGSNINKVLDYGVPENLEHITKGTAVTISLRGGKK
VGVIYQIIKTTTOCKKILPILGLESSEIVLPQDLLDLLFWISQYYFAPLGKTLKLFLPAIS
SNVIOPKGHYRVVLKQSKAKTKEILAKLEVLHPSGGAVLKILLQHASPPGLSSLMETAKV
SQSPIHSLEKIGILDIVDAQLELGDEDLLTFFPPAPKDLHPEQGSAIDKIFSSLKTSGFH
THLLFGITGSGKTEIYLRATSEALKQGKSTILLVPEIALTVOTVSLFKARFGKDVGVLHH
KLSDSDKSRTWRQASEGSLRILIGPRSALFCPMKNIGGIIVDEEHDPAYKOTESPPCYHA
KLSDSDKSRTWRQASEGSLRILIGPRSALFCPMKNIGGIIVDEEHDPAYKOTESPPCYHA
KLSDSDKSRTWRQASEGSLRILIGPRSALFCPMKNIGGIIVDEHDPAYKOTESPPCYHA
KLSDSDKSRTWRQASEGSLRILIGPRSALFCPMKPILGIIVDEHDPAYKOTESPPCYHA
KLSDSDKSRTWRQASEGSLRILIGPRSALFCPMKNIGGIIVDGEHDPAYKOTESPPCYHA
KLSDSDKSRTWRQASEGSLRILIGPSALFCPMVLIFFNRRGYHTMVSGTVCKHTLKCPHCDMVLT
FHKYAM7LLCHLCNSSPKDLFQCPKCLGTMTLQYRGSGTEKIEKILQIFFOIRTIRID
SDYTKFKGSHETLLRQFANDKADVLIGTOMIAKGMNFSAVTLAVILMGDSGLYIPDFRAS
EQVPOLITOVANGRGRENIHAVEILIQGELPDHPTTHSAMPQDYSAFYDQEITGRELCEYP
EPIRLIPGIFMGKCPKGTWEEAHBVHNILKEQLESTNPLMFYTPCGHFKIKDTFRYQFLI
VDAYYIPYNKKLHHALMLAKLDPWKFMIDVDPMTTFF

CPO_0925 TOS7915 TOS7926

CP779 hypother ical profess
RHMLEMENSONEHDTLCOLLDRYGEDLYCTLAGLENVTLPHTATGAGYGSTPEKAVEVPN
REPORTEPPPFTHLGSEKTRIGHARCVPLHFDLGONATEKEKYPALYCGSTPARKTPGST
PYVEDINEEVLEFNRLAKTLTCGLEPPKIGTLHAKTNIFYNNPNFFLALADLNYFRYKTD
TTDYHGGLTONGCTFLPLYGGLEYEKDGOLKRNLMATLNRLPFAYTPYGS 1057915 1057226

1058000 1058557 CPn_0025 1058000 1058 Thioredoxin Disulfide Isomerase CHITOTY/TREFKDSDMKFWLOGCAFVOCLLLTLPCCA. KASGENLQOTRPIAAANL OWESYABLEHSKOOMKPTCLEFTTSSDMCMMCIKNODQILQSSEFKHFAGVHLHMVEVDF PQKNHQPEEQRQKNQELKAQYKVTGFPELVFIDABGKQLARMGFEPGGGAAYVSKVKSAL

1050918 1058670 לנף לברת_חיום

VDRLAAVISNYLDSENRKSQQPD

CPn_0928 1061035 1059884

*CHLPS 43 kDa protein homolog_3

RRKDFAFTILNISNRSDILSGIFSNPHPVSYFSSTHAKQLSDFSKKHPILTKIVTIIVKI
FKLLIGLIIPPLGIYWLCQLVCSLALFPRSSMLYSVLKTCFKKYRLEQEIQDYFVKNLDP
SFKDPAVSESKRITIQOPHLTIDTLAIHFSTAPFKRWLLISLGSGDFLEDMIGLKDSLFL
SWKELAKLLGANILIYNYPGVKSSTOKLNLENLATAHNLCAKYLQDKIQCPGANEIITYG
YSLGGVVOSAALQKNPFTNSETSWVAVKDRAPHSLPAANSFFGPIGKLIAVLARWMMDA
EKNSRELPCPEILVYSADRFRPSEVGDDTALLPEFTLAHAIKRTPFARSKKFIGEVNLLH
SSPLKHPTIQKLAEAILESLSRKN

CPn_0929 1062301 1061186

*CHLPS 43 kDa protein homolog_4

KEKHAPIHGSNAPVEDILHSHPSPQATYFSSTRAQKLHEFKDRHPVLTRIASVIIKIFKV
LIGLIILPLGIYHLCOTLCTNSILESKULLKIFKQPNTKTLKTNYLHALQDYSSKURVA
SMRRVPILQDNVLIDTLEICLSQAPTNRWALISLGSDCSLEĖIACKEIFDSWQRFAKLIG
ANILVYNYPGVMSSTGSSSLKDLASAHNICTRYLKDKEQGPGAKEIITYGYSLGGLIQAE
ALROQKIVANDDTTWIAVKDRCPLFISPEGFHSCRRIGKLVARLFGWGTKAVERSQDLPC
LEIFLYPTDSLRRSTVRQNKLLAPELTLAHAIKNSPYVQNKEFIEVRLSSDIDPIDSKTR VALATPILKKLS

CPn_0930 1062851 1063330
No robust homolog present in Genebank/EMBL as of 11/7/98
NKWSELAPCSTGLOMYPHTOVHHALDTRRVILTIAACLSLIAGIVLVGLGAAAILPSLFG
VIGMULILFSSTALTYLYKKTREVDOIALEPLPENISKOQSIIDFVKTRDYASLEKKAT FAYTHTHYYDGSMVFYREIPRFMLGSYLALRKDMDRQALF

CPGT931 1064078 1065718

1955-Lysy1 trna Synchetase
1DFRYLGWKSDIYTNILEERMTARAEYLDHEDFLYRSHKLQELSELGVVLYPYEFPGVFS
1DFRYLGWKSDIYTNILEERMTARAEYLDHEDFLYRSHKLQELSELGVVLYPYEFPGVFS
CEDIKKTFASQELGNSEAAMSRSTPRVFBGRLVLFRAMGKNAFGGILDHNOTIQVMFNR
EFTSYHGLSEDAEITPIKFIEKKLDLGDIGIDGYLFFTHSGELIVLUETVTLLCKSLLS
LPPKHAGLSDKEVRYRKRWLDLISSREVSDTFVKRSYIIKLIRNYMDAHGFLEVETPILQ
NIYGGAEAKPFTTTMEALHSEHFLRISLEIALKKLIVGGAPRIYELGKVFRNEGIDRTHN
PEFTMIEAYAAYMDYKEVMVFVENLVEHLVRAVNHDNTSLVYSYWKHGQEVDFKAFWJR
MTMKESIATVAGIUDVHSDOKLKEILKKKTTFPETAFATASRGMLIAALFDELVSDMLI
APHHITOHVETTPICKTLRSGDTAFVERFESFCLGKELCNAYSELNDPIRQRELLEQOH
TKKEELPDSECHPIDEEFLEALCQGMPPAGGFGIGVDRLVMILTNAASIRDVLYFPVMRR
FDAEKTN

CPM_m0932 1067160 1065721

CYSS-Cysteinyl trna Synthetase
VKSD*VMAFSHIEGLYFYNTASOKKELFFPNTPVRLYTCGPTVYDYAHIGNFRTYVFEN
VKSD*VMAFSHIEGLYFYNTASOKKELFFPNTPVRLYTCGPTVYDYAHIGNFRTYVFEN
VKSD*VMAFSHIEGLYFYNTASOKKELFFPNTPVRLYTCGPTVYDYAHIGNFRTYVFEN
VKSD*VMAFSHIEGLYFFTQASKNIPPNTPGKLSHLDLSSLRCCZR
ISADEYDKENPSDFVLWKAYNPERDGVIYWESPFGKGRPGWHLECSIMABELLGDSLDIH
AGGVTNIFPHENELIAGSEALSGKPFARYWLHSEHLLIDGKKMSKSLGNFLTLRDLHQE
FTCGEVRYMLLQSHYRTOLNFTEEALLACRHALRRLKGFVSRLEGVDLPGESPLPRTLDS
SSOPHEAFSRALANDLNVSTGFASLFDFVHEINTLIDGGHFSKADSLYILDTLKKDTUL
GVLPHTTSVCIPETVMQLVAEREEARKTKNWAMADTLRDEILAAGFLVEDSKSGFKVKPL

CPn_0933 1067532 1068578

predicted disulfide bond isomerase

PVILLONIRRCSLKQLKVLATLLLSLSIPTLEAAENRDSDSIVWHLDYQEALOKSKEAEL

PLUVIFSGSDWOPCHKIRKEVLESSEPTIKRVQKFVCVEVEYLKHRROVENIRQONLAL

KSKFKINELPCMILLSHEEREIYRIGSFONETGSNLGDSLCHIVESDSLLKRAFFMWTSL

SLSELDGRYYRLAEELSHKEFLKHALELGVRSDDYFFLSEKFRLLVEVGKYDSEECQRIKK

RLUNKDPKNEKQTHFTVALIEFOGLAKRSRAGVRODASOVIAPLESYISQTGQODKDNLW

RVEMMIAQFYLDSDQWHHALQHAEVAFEAAPNEVRSHISRSLEYIRHO

1068948 1068526 cri_0534
rnpA-Ribonuclease P Protein Component
rrvhPLTLPKQSRVLKRKOFLYITRSGFCCRGSQATF/VVPSRHPGTCRMGITVSKKFGK
AHERNSFKRVVREVFRHVRHQLPNCQIVVFPKGHKQRPVFSKLLQDFINQIPEGLHRLGK TKATTGGECTPKSEKCVTAPR

1069100 1068957 CPn_0935 1069100 1068957 r134-L34 Ribosomal Protein EDTVKRTYQPSKRKRRNSVGFRTRMATRNGRKLLNREPRHG RHSLVDL

1069330 1069470 YLMKVOSSVKADPSKGDKLVRRKGRLYVINKKDPNPYSRQAGPARKK

1069487 1069798 TOTA - 014 - RIEXERONMAL PROTEIN
VKRMAKKENVARFAKRREVEANFKKREDLRKIVKELDVEEEKENARISLNKMKRDTSP
TPLINNECLITORERGYLRKFALDRICFROMADMOVIEVYLKADW

1070355 106984 CPTG_0038 1070345 106984/ CPTBB hyperhetical protein -flagfer (60) peptide periplosmicl CPTBBCYTERFMEHTICLEFYVILOCUSAY (ØDKKYFITY ITWEFASAFFOFTGLOVULLL EGRRIALERGYTERFFNORLFODLKKOLACHOETGOODLOETY LOTEKWEYLNKORENV CHED BEEFFAATTIK GRAA BEET MAMKKOWKD MÕKAKDA SOOOTIKEVSK

1071195 CPR_1933
CT730 hypothetical page
HINRWTIRLSLTLITSTVLYFFSEEIELLOGGKMEKQNLKLDVKEIEFPETVFSRDIETR
VIOVIILLCLAKINGVSLLOGGKLIDGLEGRDIERNKGIYVEQDSKNHLVKVRVEVNVDYG
VSIPEKTEEIQOCIVSEISEYTGLHVAAVHVIIKGLTQPKDRIDEEIEEEVSVQDLPSPE DFLLENSEC

1071204 1073033 CPn .0940 DATTE LUMPHERE LIVER OF VYSME PROPOSITE ARE SEPLAT SHEKIDUR ERIPFLMKKTAS IET IVVSNET PALLLENNLIKOHHPKYNVLLKDKTFFCLAISLSHSW PKVEAIRTKAITSSORQLIFG FYVSAEACHTLLEVISOM FPLRTCSDREFALRKRE ILY DMKRCLARCVGYCTPEEYOGT DKAILFLKCKIEEVVKDLEKVIQKASDNLEFEQAANLY RTLSLIKOAMAKOQVEKPHFØNIDALGLYRHKORTILTLLTVRSCKLLGARHFSFF ENAO EDODLLSSF ILOYVYSOPY FREILTPLPLEF PTLSYVLNAESPRELRSPKTGYGKELLD LAYRNAKAYAATTLPSSTUPYODFON ILRMSQYPYRIECYDNAHMOGAHATCYYIVFEN GFDPKOYRTFSIDSEKTOMDLALLEEVLLRREHSLTTALPOMIVVDGGKTHYNKTKKIN TILNLTGIOVVTIAKEKSMHSRGLNKEKIFCETFPEGFSLPPTSNLLOFFOILADEAHRFA ISKHRKKRGKALFEQEYIPGIGEVKRKRLLOKFKSWKOVMLSSQEELEAIPGLTKKDIAV LLAROKDFNKSD Toyet Excinations DATES ILVMRIEDECTEL arrani. ANATWANAMANAN CAMPOLINABATA ANDREKODOR

CPN_0941

mutS-DNA Mismatch Repair

VMTEKKPTPMMEOMHOCKEKAGDSVLLFRMGDFYEAFYDDAVLLSOHLELTLTORQGIPM
SGIPVSTVDTYVDRLIGKGFKVAVAEOFGEPAKEKESKKIGPMARDIQRFVTFOTLLSST
LLOEKFNNYIVAINRIGSLFGFACLDLSTGSFFIEECENTKELVDEICRLAPSEVLSONL
KYNKETALVMGLOOHLKUTLSTYADMAFEHKFASOKLTHTFOVASLDGFGLKGLVPAINA
AGGLLSYIODKLLLPTKHIAIPOTRGKOOKLLIDTASOVNLELLAPLNDPOGKNSLLRIM
DHTSTPMGGRLLRGILISPFYNPKEILVRODAVEFFIROVILRRIKTYLQOVRDIERLI
KTVTTGLAGPRDIGTLRDSFSAGAQIYEOLASATLPEFFIDKCSLDTKLASLIALLSKSL
NGDLPLRVSDGNIFVDEFHNDLKRLRHNGEHSGEMIWEYQERIKKETGIKKKKIGFAQAL
GYYIEVSEFFAPOLPKOPIRROSGIHLAERFTTIELOOFODDMSNISEKLOTLETOFFKOL
CYYIEVSEFAPOLPKOPIRROSGIHAERFTTIELOOFODDMSNISEKLOTLETOFFKOL
CYYIEVSEFAPOLPKOFIRROSGIHAERFTTIELOOFODDMSNISEKLOTLETOFFKOL
CYTIEVSKEFAPOLPKORGNITHVEMARGTYPIRQIALLVIMAOMGSYIPAKSAHIO
UDKIFTRIGAGDNISKGKSTTMVEMAETANILMATDRSLVILDEVGRGTSTYDGLAIAQ
AVVEZLLFTDKKKAKTLFATHYKELTTLEDHCPHVENFHAGVKDKAGQPVFLYEILKGHS
QKSYGIHVARLAGFPLCVVSRAQQILROLEGPESITRPAQDKMQQLTLF 1075504 1073018

1075955 1077754 CPÁ_0942 1075955 1077754

d/ad/prim-dna Primase
wcsitklrtamyteesldnlrhsidivdvlsehihlkrsgatykaccpphtektpsfivn
bagahyhefoggahgdaigflmohlgysfteailvlskkfovdlvlopkdsgytppoglk
belrhinsbeatfffryctyhlpearhaloytyhrofspdtidrhlgygeoslfloame
erkisoeolhtagffonwflfarriifpyhdalchtiofsarkflensoggkyvntpet
pifkksrilfglnfsrriakekkvilveogadclomidsgenctvaagotafteehvel
sklgvlkvfllfdsdeagnkaalrvgdlcotaomsvfvcklpoghdddsflworgsgl
lalleosodyltflisemssypkfoprekallveeairoikhwgspilvyehlkolasl
mwypedmylslanpoytaeponipilkokypkthphivmetdilcchlfosntkiltao
fyfypedfkhecrklfafmisyyekyrknvpfdeacovlsdsoilolltkrrinteald
tifvoslokmadrkwreockplslmoniodkkleiledyvolrkdrtittlldpeselip CPA_0942

1077972 1078238 CT794.1 hypothetical protein
FFMKSFKFLLPFLSVILCCGNLLSSPRSRAISVTESIGMSAVKTLVLSEKAHEFLEGIGY
GVGASSILRDWQTQQWLEIESLLAQNEVM

1078503 1078997 CPn 0944 CPI_0944 1078503 1079997
No robust homolog present in Genebank/EMBL as of 11/7/98
IKIMMRYFIPLLALLIFSPSLVRAELOPSENRKGGWPTQLSCAEGSQLFCKFEAAYNNA
IEEGKRGILVFFSERPTPEFADLTNGSFSLSTPIAKGFNVVVLCPGLISPLDFFHKMDPV ILYMGSFLEMFPEVEAVSGPRLCYILIDEQGGAQCQAVLPLETKN

CPn_0945 1079001 1079660
CT795 hypothetical protein
SIFRMKILPSYFGHNFDQLRRHYMRIALSLLSLLMIFPIFGEESRPGSEDGNSNTQEIVG
SQDTQVCLYHSYEQGLQASRIEGKPLVIVVLCNSGDDQQACTIGLSETCEEVLSVLSGSI
FSELANFVVLVPSGVNPLIYPPIEDPILAEIVKFKELFKDESFPTGLSIIVVGVTPEGPG
DIIEVSPVSLTVEEEETLPSEQTTEVESTSELQSEDPAIA

CPn_0946 1082816 1079745
glyO-GlyCyl tRNA Synthetase
GCQKKKCYTLESFVSEHPLTLQSMIATILRFWSEQGCVIHQGYDLEVGAGTFNPATFLR
ALGPEPYKAAYVEPSRRPQDGRYGVHPNRLQNYHQLQVILKPVPENFLSLYTESLRAIGL
DLRDHDIRFIHDDWENPTIGAWGLGWEWWLNGMEITQLTYFQAIGSKPLDTISGEITYGL
ERIAWYLQKKISIYDAYDEVIKQO ITQASEKAWSENNFDYANTEMMFKHFEDFAEFAL
RTLKNGLSVPAYDFVIKASHAFNILDARGTISVTERTRYIARIRQLTRLVADSYVEWRAS
LNYPLLSLSSTSEPKETSESVVPMISSTEDLLLEIGSEELPATFYPIGIQDLESLARQVL
TDHNIVYBGLEVLGSPRRLALLVKNVAPEVVQKAFEKKGPMLTSLFSPDGDVSPQGQFF
ASOGVDISHYQDLGSHRASLAIRTNVGSEYLFLLHEEIRLRTADILMQELPLLIGPMKFPK
KMVMDMSQVEYARPIRWUALYGEHILDFITGTITASRNSFGHRQLDPRKISISSPQDVV
ETLRQACVVVSQKERRMIIEQGLRAHSSDTISAIPLPRLIEEATFLSEHPFVSCGQFSEQ
FCALPKELLIAEMVHNGYFTHETSSGAISNFFIVVCDNSPNDTIIEQNEKALTPRLID
AIQYCKADLVSAVVNEFPELQGIMGEYYLKHANLETASAVAVGEHLRHITMQGKLSTIGT
LUSLLGRLDNLLACFILGKFTSSHDPYALRRQSLEVTLVGASRLPIDLGSLGDLDAH
PSTIEEKVWDKSKTIHEILEFIWGRLKTFMGSLEFRKDEIAAVLIDSATKNPIETLDTA
EALQLLKEEHTEKLAVITTTHNRLKKLLSSLKLSMTSSPIEVLGDRESNFKQVLDAFFGF
PKETCAHAFLEYFLSLADLSNDIQDFLHTVHIANDDGAIRNLRISLLLTAMDKFSLCHWE
SVAV 1082816 1079745

1083433 1084059 GENERAL CONTROL 1085483 1684047 SSEAARETOKOOGRAFIAARATTITUTOTOTETELAAATTITUTATATTITUTATTITUTATATTITUTATT

YLQEADPADIYHLHDWHYGLLAGLLKNPLNPVHSKIVE DPHLGHYOLFRDFOTSVLMKGALYCSDYITTVSLTYV GGIINGIDEDVWNPKTDPALAVOYDAGLLGEPDVLFTK SDYELHDA ILARNSVE DFHLSHYOLFROLVTSVLMKGALVGSDYTTTVSLTYVL
SGIINGIDEDVMNPKTDPALAVQYDASLLSEPDVLFTKN AVVLYEKLGISSDYFPLI
CVISRIVEEKQPEFMKEIILHAMEHSYAFILIGTSQNEVLLNEFRNLQDCLASSPNIRLI
LDFNDPLARLTYAAADMICIPSHREACGLTQLIAMRYGTVPLVRKTGGLADTVIPGVNGF
TFFDTNNFNEFRAMLSNAVTTYRQEPDVWLNLIESGMLRASGLDAMAKHYVNLYQSLLS

1085887 1086483

AGIETITPLKEVAITVSRR

CPn_0950 1086470 1087027
pth-Peptidyl trna Hydrolase
PSLEDMMAKLIVAIGNPRHGYANTRHNAGFLLADRLVEELQGPPFKPLSKCHALMTLVES
SSGPLVFIKPTTFVNLSGKAVVLAKKYFNVALSHILVLADDVNRSFGKLRLCFNGGSGGH
NGLKSITASLGSNEYWQLRFGVGRPLEBGVELSNFVLGKFSEEENLQLGSIFVEASTLFT

CPn_0951 1087113 1087457
rs6-S6 Ribosomal Protein
eFLMGKKENQLYBGAYVFSVTLSEEARRKALDKVISGITNYGGEIHKIHDQGRKKLAYTI
RGAREGYYYFIYFSVSPGAITELWKEYHLNEDLLRFMTLRADSVKEVLEFASLPE

CPn_0952 1087469 1087723
rs18-s18 Ribosomal Protein
GEMMNKPVHNNEHRRKRFNKKCPFVSAGWKTIDYKDVETLKKFITERGKVLPRRITGVSS
RFQGVLSQAIKRARHLGLLPFVGED

1087727 1088248

CPI_0953
T19-L9 Ribosomal Protein
FKGRMKQOLLLLEDVDGLGRSGDLITARPGYVRNYLIPKKKAVIAGAGTLRLQAKLKEQ
RLIQAADKADSERIAQALKDIVLEFQVRVDPDNNMYGSVTIADIIAEAAKKNIFLVRKN
FPHAHYAIKNLCKKNIPLKLKEEVTATLLVEVTSDNEYVTVLAQGKQTEENQEG

CPn_0954 1088259 1088708
ychB-Predicted Kinase
GRKVCYKDIMQYFSPAKINLFLKIWGKRFDNFHELTTLYQAIDFGDTLSLKNSMKDSLSS
NAMELLSPSNLIWKSLEIFRRETQIHQPVSWHLNKSIPLQSGLGGGSSNAATALYALNEH
FÖRHIPITTLQLWAREIGSDVPFFFLQEQH

CPN-0955 1088612 1089175 CFTEU993 (IfZme-shift with 0954) RAFFNPYSYNNIATLGSRNRKRGSFFFSSGTALGKGRGEHLFSIKKLNHKHKYVLYLDHQ GEMMERAYOSLLPODYSTONNACFYGENDLEKSVFRIRTDLKNKKHMLERMWSFFESHV LMSGSGATLFVCYLEELEQDSKVSSQIHSLIKQTQGIPVSRLYREPHWYSLKQSTYKNSP

CPn_0956 1089545 1090909 CHÉ-0956 1089545 1090909

CT865 hypothetical protein
LVWÉSMILPPYSYSLKIGAAVLFFCSILHTFLTPMLYTLCQSYEHKKLVFPECWKRYARL
SEHERILSRVEIVFFLWAVPLFFWFLYTEGYRISWAYFNSRNYGFAVFIMVILILLESRP
IVWFAELVLSSIAKLGKTSPKSHWNTHLTAPPLLSCLLKETGAMIGATLHRAFFYVFSP
SRFAYATMGLLFSNISIGGLTSYVSSRALFLIFPALKWEHSFFLSHFAWKAIVAILIST
TTYYFIFRKEFKKFPDIPSDKDPSVEKVPWNIICVNIIFVGSIILSRSTPLFMGALLLFY
LEPEKFFIFYODPINLSKVCYVGLFYAGLVVFGLOEWMVLINHGGLSDFGYMTVSYTLS
IFILDNALVNYLVHNLSVATDCYHYLVVAGCMAAGGLTLVSNIPNIVGYLILRSAFPSSTI
HWEMLFLGALGPSIISLGVFWLLKNVPEFLYCFFR

THE OPST 1093812 1090963

ide/ptr-Insulinase family/Protease III

KIMTRNCKMFWKLLCPILLCTSLSITSCEOOFKVVPNCPLOVSTPAAADOKIEKIICSN
GLEBLIISDPNLPTSGAALLVKTGKNADPEEYFGMAHFTEHCVFLGNEKYEVSGFFGFG
SENGVKNAFTYPNKTVFVFSVEHSAFSDALDOFVHLFINPKFRQEDLDREKYAVHQEFA
AHBESDGRWHRIOQULVAPOGHECAFFGGNASTLTPVTTEKMAEWFKLHYSPENMCAIA
YTSAPLSKAKKQFSKIFSQIPRSKNYERQEPFLPSGDTSSLKNLYINQAIQPTSNLEIYV
HIVESSHPIPLGCYKALAEVLRNESKNSLVSLLKNEQLITOLDVEFFRSSLNTGEFYISY
ELTEKGGKHYSQVIDSTFOYLRYIQEHGIPNYTLEISTINALYCYSSKSPLFDLLCKQ
IVSLGNEDLSTYPYHSLVYPKYSSEDESALLNLVSDPEOARFVLSSKNSEHWEEATOWHD
BIEDMTVAKAL DOWNOWLOKSUK BLALDKDNIF IPMFVTLIGVHLLKKGEFFPABALS IVSIGNEDLSTYPYHSLVYPKYSSEDESALLNLVSDPEQARFVLSSKNSEHWEEATOCHD
PIFDMTYVYKALDGVODYGKVOSLKPIALPKPNILIPKEVTLEGVHLLKKGEFFERJALS
YQDDKLTLYHCEDHYYTAPKLSSQIRIRSPQISRSSPQFLVATELYCLAVNDQLLREYYP
ATOAGLSFTSALGGGIDLDRYGSYTTIVPALLNSILTSLPNLEISYEFFLYYKKQLLELY
GALLNCPYPRSGLDELASQVMKETYSNTTKLSALEKLSFSEFGAFASNLFNSVHLEVMYL
GNLSEQOKKDYLEMLQVFTASRSSHATKFPYVELQSQEISEIHHDYPLTANGML/LLQDK
SSPSIGKVCAEMLFEWLHHITFEELRTQQLCYMVGARYREFASRPFGFLYIKSDAYSP
EELLAATTSLFLNKVSASPEKFGISQEKFANIRKAYINKILEPEHSLDMMNSAUFSLAFER
PFVEFSTPDLKIAIAETLTYEEFLKYCQCFLSNELGTQTSVYIRGTQKTS

1094803 1093793 CPn 0958 CPH_0958 1094803 1093793
plsB-Glycerol-3-P Acyltransferase
tyraiymofsrylryafdnoylpeplyokfsvfhonyidaatkkaaadolevlclowkv
iiedlknipfifppyhkkirapiolffilsidffslviddknskilalhrukeieeyiargd
nvvllanhofecopolmyyalgktheelmennifvagdrvtsdplarpfsmcolllciys
krhiatppelreekllhnoksmoilktillneggkfiyvapaggrorkmaegrlypsefsp
esievfrllakasnotthfypfalktydilppppkienaigegraiyfapvffnfgaelf
fdalcokeelihcdkhaortlraekvfsivknlyeel

CPG_0959 1096376 1094799
GatE-Azial Filament Protein
ACGYGIGTRKVMENEILLNIESKEIRYAHLKNGQLFDLTIERKKVRQLKGNIYRGRVTNI AQLYGITTRKWMENTLLINIESKETRYAHLKNÖQLEDLTTERKYVRÖLKGNIYAGRYVNI LARNIGUAFINI DERENGISTH HIGOTLE DENGKFEQNEPOMOVDAL PERASCAPILISSERAFIE EFIKLDUPVI.VOVVKETIGGKGARLTSNISTEGRYLVLLPNOPHRGVSRKTEDPHMREQL KQLI RUPERHEJOM BLI CREASTEAUTEAL I DENGHOLLUWKY TLEKFYSTEQECLUKSET DILKKYN FORTOL DENKYNTOKCKHINILKKYSPOASIK LEYYRDST PMEERNIE REI DIKAPHIKI KWOUGANYLEFIOKTEAMIET LOVINGOSTOJUGGOVETLVOI NILEABETA ROLKHINIVALIVI TJOFI LOMKORKNORRVLERLKEHMKYDAARCT I LOMSEFGLVÆTTROR NREDLAUPTUFTLOPYCHANAT I KTDESVVIET LERDLING I NIKKEHSHUCLVVHPETASYM KÖRHUDHEMINLAKOLKAKLO INTODOVHUNHYOFFSUTTGES I DL

CPL_0960 TOTTOS
CT809 hypotherical Control Con

1097106 10/17277 rl32-L32 Ribosomal Profein RVCCMAVERSE TEACTURE RIPARE SECUTION OF STREET

CPn_0962 1997301 1098275
plsx-FA/Phospholipid Synyhesis Protein
ILSOPMEVQIGIDLMCGDHSPLVMCVLVDVLKSQSSTIPFAFTLFASEEIRKQIQEEFI
SDLPQEKFPKIISAENFVAMEDSPLAAIRKKSSSMALGLDYLQEDKLDAFISTGNTGALV
TLARAKIPLFPAVSRPALLVCV9TMRGHAVILDVGANISVKPEEMVGFARMGLAYRQCLG
DSKIPTIGLLNIGSEERKGTEAMROTFFMLRETFGEAFLGNIESGAVFDGAADIVVTDGF
TGNIFLKTAEGVFEFLORILGDKLEADIQRRLDYTFYPGSVVCGLSKLVIKCHGKACGSS
LFHGILGSINLAQARLCKRIJSNLI

CPI_0963 1698374 1103224

pmp_21-Putative Ouyer Membrane Protein
TPLERKVANVAKKTYRSVSSFSHSVIVALISAGIAFEAHSLHSSELDLGVFNKOFEEHS
AHVEEAOTSVLKGSDPVPSCKESEKVLYTOVPLTOGSSGESLDLADANFLEHFONLFEE
TTVFGIDOKLWSDLDTRNFSOPTOEPDTSNAVSEKISSDTKENRRÜLETEDPSKKSGLK
EVSSDLPKSPETAVA/ISEDLEISENISARPLOGLAFFYKNTSSOSISEKDSSFGOILF
SGSGANSCLGFENLKAPKSGAAVYSDRDIVFENLVKGLSFISCESLEDGSAAGVNIVVTH
CGDVTLTDCATGLDKEAURLYKDFSRGGAVTTARNHEVONNLAGGILSVVGNKGAIVVEK
NSAEKSNGAFACOSFVSYNBENTALWENOALSGAISSASDIDIOGNCSAIEFSSMOS
LIALGEHIGLTDFVGGGALAAOGTLTLRNNAVVCVKNTSKTHGGAILAGTVDLNETISE
VAFKONTAALTIGALSANDKVIIANNFGELIFEONEVSHNIGGAIVCGCRSNPKLEOKDSG
ENINIIGNGA/TFLKNRASVLEWTYQAEDVAGGGALMGHNVLLDSNSGNIOFIGNIGGS
TFWIGGSVGGGALLSTDRVTISNNSGDVYFKGNKGQCLAOKXVAPOETAPVESDASSTNK
DEKSLNACSHGDHYPPKTVEEEVPPSLLEEHPVVSSTDIRGGGAILAGHIFITDNTGNLK
VDISANHSVEFVSNGSGKFGGAVCALNESVNITDNGSAVSFSKNRTRLGGAGVAAPOGSV
TICGNGGN/AFKENFVFGSEONSCGGAILANSVNIODNAGDLIFVSNSTGSYGGAIFV
GSLVASEGSNBRTLTITGNSGDILFAKNSTQTAASLSEKDSFGGAIYTONLKIVKNAGN
VSFYGNRAPSGAGVQIADGGTVCLEAFGGDILFEGNINFDSSFNSTTSYGGAIFV
GSLVASEGSNBRTLTITGNSGDILFAKNSTQTAASLSEKDSFGGAIYTONLKIVKNAGN
VSFYGNRAPSGAGVQIADGGTVCLEAFGGDILFEGNINFDSSFNSTTSTSYGGAIFV
GSLVASEGSNBRTLTITGNSGDILFAKNSTQTAASLSEKDSFGGGAIYTONLKIVKNAGN
VSFYGNRAPSGAGVQIADGGTVCLEAFGGDILFEGNINFDSSFNSTTSYGGAIFV
EETLYSAGVQINMSSPTENKOKAVDTPVLADIISITVDLSSFVPEQDGAIPTPENK
EETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGALPPEPEIINK
EETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGALPPEPEIINK
EETLYSAGVQINMSSPTENKOKAVDTFVLADIISITVDLSSFVPEQDGTLPPPEIINK
EETLYSAGVGINMSSPTENKOKAVDTSVLADFUSCHVINPERGGALVLNNINSHTYDLAR
LKOEIFAHHTIAQRNELDFSTNVWGSGIGVVEDCONIGEFGFKHHLTGYAGLGIDTOLVE
DFLIGGCFSOFFGKTESQSYKAMOVKSYMGAAYAGILAGPWLIKGAFVYGNINNDLTTD
VYDTLGISTGSNIGKGFIAGTSIDVRYIVNPRRFISAIVSTVVPFVEAEVRILLDEDESDE
CKEVNTFOKTYRPPPATAGHTPURGFALLASYRGGRAEVNSVOLAYVYDVYKRGPVSLITLKDA
AYSWKSYGVDIPCKAWKARLSNNTEWNSYLSTYLAFNYEWREDLIAYDFNSGIRIIF

CPn_0964 1104812 1103301
No robust homolog present in Genebank/EMBL as of 11/7/98
QSILESIIKYFYLIKNSKMHMSNPISLFSPAELIAKYNLIPKTSPIYPRRTELIILEENA
COTRLITNVAQVLHPSSLFSMSKKILNPCGCSGGPLCWVILNILAFIITSVLFIILLEVNL
VAGGLRIPMPLPPKKIVEDLSEPTTEETNEVIQPFIFALQALLFEDNKLRSFKIVEGSVG
KAPLPNPFLNRLVAISPQESQEAMRKIPDLCSQLKKVLKSLGVLTPEWKHMLKYFEGLKN
HHDSNPDKKTFPILIKLLIEALTGKSSLPKTFSTKEKNQAALFIASSCKTCKPTWGEVIT
KSLNRLYSIANEGONQLIWVOEFKERELMSIQDGDAEEYRFAAQQHGERYTEAIQVL
RNESAAKLQWHVINTMKFFHGKNLGLVTEHLQDTLGALTLRQTTVDTHQGREDADLSAAL
FLNRYLNSGNQLVNSVFKSMQKADPETKALIREFALDILYASLRLPQTSAHTEVFSTLLM
DEPTYFEDNKACIAVILVJKKIFI. DPETYEPNKACIAYLLYVLKIIEL

CPn_0965 1106769 1104925

1pxB-Lipid A Disaccharide Synthase
KGFSFSKVGLMMIPSGLVYLLYPLGFLASLFFGSAFSIQWMLSKKRKEVYAPRSFWILSS
IGATLMIVHGTIQSGPVTVLHVINLITYLRNLNITSSRPISFRATLVLMALSVVFVTLP
FLYVNMEMMASPNIFHLPLPPAQLSWHLIGCLGLAIFSGRFLIQWFYIESNNTKDFPLLF
WKIGLLGGLLALVYFIRIGDPINILCYGGLFPSIANLRLFYKEQRSTPYLDTHCFLSAG
EASGDILGGKLIQSIKSLYPNIRFWGVGGPAMRQBGLQPILMMEEPQVSGFAEVLGSLFR
LYRNYRKILKTILKHKPATLIFIDFPDFHLLLIKKLRKHGYRGKIIHYVCPSIWAWRPKR
KRILEQHLDMLLILPFEEGLFKNTSLETVYLGHELVEEISDYKEQASWKEKFLNSDRPI
VAAFFGSRGGISRNLRIOVQAFLMSSLSGTHQFVVSSSSAYYDEIIEDTLKAEGCOHSQ
IIPMFRYELMRSCDCALAKCGTIVLETALNGTPTIVMCRLRFFDTFLAKYIFKILLPAY
SLPNIIMNSVIFPEFIGGKKDFHPEEIATALDLLNQHGSKEKQKEDCRKLCKVMTTGQIA
SEEFLKRIFDTLPAV SEEFLKRIFDTLPAV

1108055 1106748

CPG_0966 1108055 1108748

pcnB_2-Polya Polymerase
LLITIIMVCENNILSGRGLELLKKKSNITLTFTIYSVSNHNIKLKDFSPHALSVIKTLRK
AGYTAYTVGGCIRDLLLNTTPKDFDISTSAKPEEIKAIFKNCILVGKRFRLAHIRFSKQI
IEVSTFRSGSTDEDVLITKDDHLWGTPEEDVLRRDFTINGLFYDPEHEEIIDYTGGVMDLR
NRYLRTIGDPFTRFKQDFVRMLRLLKILSRSPFTVETGTQEALIACRQELIKSSQARVFE
ELIKMLNSGAAKNFFQLLIENHLEEILFYMDKAFFRLNPALEEGTATYLKALDDKILKKE
AEYDRHQUMAIFLEPLVNFNVRYKHQKHPYLSLTSVFDYIKNFLEGFFADSFTSCSKKNF
ILTALILOMGYRLTPLIPTKKALFFNKKLLHHTRFLEALSLLEIRSIVYPKLDKVYVAWI RHHOTLKCKKDSHSQK

1108431 1109885 CPn 0967 CPG_0967 1109431 110985
mrsA/pgm-Phosphositucomuciose
FTAYKFAFICACROEKIRRIGIDFRRNMQSVRKLFGTDJVRGRANFEPMTVETTVLLGK
AVARVLREGR:SKIHKVVGKDTRESGYMFENALIAGLISMSIETIJVLGFIFTROVAFITR
AYRADASIMISASIHIPYNDROKIK IF ISLEGFKIGIJVLEGFIETMVSEADPROFIDEDBHAVGK
NKRVIDAMGRYVEFVKATFIRGETIJKOLKIVLEGAHGACYKVAFSVEEFIJAEVICYGGE
PTGININERGALFIQVIQNAVIENGAHLGIALDJEDDE I IMVDEKGHIVDISMILSICA
GDLKRRJALPHINKVANTINTMFGVIJKYLEGALIAJVETOSYGORHVIJIAMLEHEVTIJVSEQ
SCHMIFLOVNTTEN GUIVEALAVIRIM FERESMIJDJETAFIVKSHQVILINVAVREKITLEET
GELFERER HAVADALIAJALIKH MIERIESMIJDJETAFIVKSHQVILINVAVREKITLEET
GELFERER HAVADALIAJALIKH MIERIESMIJDJETAFIVKSHQVILINVAVREKITLEET CPLTERTERDVQDALGESGRELLERYSGTENTGRAMVEGHKKHQVDCLAKALADVTDAELG TGGRE

1104889 CPn_0968 1109889 1111721 qlm:-Glucosamine Fruetone o P Aminotransterane

DRMCC: FFCYLGNQDGVS: IVLEGLAKLEYRGYDSAGLAA QERE: FTASV: IGHTRWATHGVPTE: INAHPHVDEGRSCA GFASDTDSE: IVQLFSLYYQESQDLVFSFCQTLAQLRGS FIRKTVGRVOELSNLF IENFKELRRELTAQGI LIHKDHPHTILCASQES PLILGLCKEETFIASDSRAFFKYTRHSQALASGEFAIVSGGKEPEVYNLELKKIHKDVRQ ITCSEDASDKSGYGYYMLKEIYDQPEVLEGLIQKHMDEEGHILSEFLSDVPIKSFKEITI ITCSEDASDRSGGYYMLKEITVQFEVLEGLIQKHMDEEGHILSEFLSDVPIKSFKEITI VACGSSYHAQYLAKYI IESLVSTPVHIEVASEFRYRRPYIGKDTUGILISOSGETADTLA ALKELRRRNIAYLLGICNVPESAIALGVDHCLFLEAGVEIGVATTKAFTSQLLLLVFLGL KLANVHGALTHAEDGSFTOGLOSLPDLGGKLLANESLHSWAQPYSYEDKFLFLGRRLMYP YZMEAARIAREHAYTA YARAMARIAKEITA YORUMAYAMALASAMEIDG DANAMARIA LARHEDGAAVAGALASAMEIDG PRNLAKSVTVE

CPn_0969 1111803 1112999 tyrP_1-Tyrosine Transport_1 VYVMSNKVLGGSLLIAGSAIGAGVLAVPVLTAKGGFFPATFLYIVSWLFSMASGLCLLEV VYYMSNKVLGSSLLIAGSALGAGVLAVPVLTAKGGFFPATFLY IVSWLFSMASGLELLEV
MTWMKESKNPVNNLSMAESILGHVGKISICLVYLFLFYSLLIAYFCEGGNILGRVFNCON
LGISWIRHLGPLGFAILMGPIIMAGTKVIDYCNRFFMFGLTVAFGIFCALGFLKIQPSFL
VRSSWLTTINAFPVFFLAFGFGSIIPTLYYYMDKXVGDVKKAILIGTLIPLVLYYLMEVV
VLGAVSLPILSQAKIGGYTAVEALKQAHRSWAFYLIAGELFGFFALVSSFVGVALGVMFDL
ADGLKNNKKSHPFSIFFLTFIIPLAWAVCYPEIVLTCLKYAGGFGAAVIIGVFPTLIVWK
GRYGKQHHREKQLVPGGKFALFLMFLLIVINVVSIYHEL

1113452 1114648

tyrP_2-Tyrosine Transport_2
VYVMSNKVLGGSLLIAGSAIGAGVLAVPVLTAKGGFFPATFLYIVSWLFSMASGLCLLEV NTWIKESKNIPVINLISMAESILGHVÖKISICLVYLFLFYSLLIAYFCEGÖNILGRVFNOG LGISWIRHLÖPLÖFAILMÖPIIMACTKVIDYCNRFMFGLTVAFGIFCALGFLKIDPSFL VRSSWLTTINAFFVFFLAFGFÖSIPTLYYYDKKVGDVKKAILIGTLIPLVLYVLWEVV VLGAVSLPILSOAKIGGYTAVEALKÖAHRSWAFYIAGELFGFFALVSSFVGVALGVMÖFL ADGLKWNKKSHPFSIFFLTFIIPLAWAVCYPEIVLTCLKYAGGFGAAVIIGVFPTLIVWK GRYGKQHHREKQLVPGGKFALFLMFLLIVINVVSIYHEL

1114693 1115415

yccA-Transport Permease SCHOOL FEMILISOF FEMILISOF FOR THE STATE OF

CPG_0972 1116377 1115430

ft\$\(\frac{1}{2}\)LCOLOR DIVISION PROTEIN FLSY

RCIRNSLLFPSYLVSFLLLQLILLLAMFKFFRNKLQSLFKKNISLDLIEDAESLFYEADF

GTELFEELCARLRATKADASTIKDLITVILRESLEGLFSQASQSSQTRFIVSLLLGTNG

SGK\(\frac{1}{2}\)GTAKLAHYYKERSESVMLVATDTFRAGMDQARLMANELGCGFVSGQPGDAAAI

AF\(\frac{1}{2}\)GSAIARGYSRVIIDTSGRLHVHGNLMKELSKIVSVCGKALEGAPHEIFMTVDSTL

GN\(\frac{1}{2}\)EQUVRVFHDVVELSGLIFTKVDGSAKGGTLFQIAKRLKIPTKFIGYGESLKDLNEF

DL\(\frac{1}{2}\)FLINKLFPEVEKI

CPT_0973 1116346 1117527

sicG-Succinyl-CoA Synthetase, Beta
EGKSKELFMHLHEYQAKDLLASYDVPIPPYWVYSSEEEGELLITKSGLDSAVVKVQVHAG
GRÖKHGGVIVAKSSAGILQAVAKLLGMHFTSNQTADGFLPVEKVLISPLVAIQREYYVAV
IMDRWHRCPVLMLSKAGGMDIEEVAHSSPEQILTLPPLTSYGHIYSYQLRQATKFWHEWGE
VMHGGVQLIKKLAKCFYENDVSLLEINPLVTLLEGELLVLDSKITIDDNALYRHPNLEVL
YDPSQENVRDVLAKQIGLSYIALSGNIGGIVNGAGLAMSTLDILKLHGGNAANFLDVGGG ASOKOJOEAVSLVLSDESVKVLFINIFGGIMDCSVVASGLVAVMETRDQVVPTVIRLEC NVE<u>IG</u>KEJVQQSGJPCQFVSSMEEGARRAVELSM

1117523 1118422

CPR_49/4

1117523 1118422

"sufeD-Succiny1-CoA Synthetase. Alpha"

VCRERRYMFHSLSKNTPIITOGITGKAGSFHTEQCLAYGTNFVOGVTPGKGGTLWLDLPV

YDSVLEAKCATCCRATMIFVPPYAAEAILEAEEAGIELIVCITEGIEVRDMLEVARVMD

NSTBÖLIGPNCPGIIKPGECKIGIMPGYTHLENGTWYSRSGTLTYPEAVWGLTDQKICOS

ICVGLIGGDPLNGTSFIDVLQALEEDPYTELILMIGEIGGSAEEEAAAWIQAHCYKPWAF IACATAPKGKRMGHAGAIISGNSGDAKSKIQVLRESGVTVVESPAHIGKTVDAVLRAKEL

1119038 1119637

No robust homolog present in Genebank/EMBL as of 1/7/98
GIEEQVALSIAIRLKIILAILLFPLVLLAWVIRYQLHANFKCSVVPFPOFSVNQAYKCS
EAKIEEMLOLLDLETLEWSSRCLRODMTFANRLEEELIOELRVSETEEVISLGGKRNLVR
LLLTHFFNPPKRSRVESVGHEVVFPVFDRLKREEELIGGGPTTRSNEELWALLDHGTARG IHKTLWFSIFFKYLTQIELF

CPL_0976 1120079 1121185
No robust homolog present in Genebank/EMBL as of 11/7/98
LUMLVYCFDPSVPTSPEHRLMAALDRWFFLGGHRAFILTHEGNHYRAFGENNSISTVEKI
LKLISYLLIFIVLIALLIRCFHLSRFKCMWKCDSLSDARVPHDVOFFNDFOLFNNOERLN
IWKNRRVVSGIDVLMVPVDYLRSOFFDFKE1PEAIRCENVVSDGOFSEESKTSYLRAMLT
DIVGYILSLDETYWTNVILKIRAMCITFESFFGKEADDNYSBRVTHHYFDESWKALARHV
LGEGNHVNRLDEALIRTERFGKEGECITKQFLKDYCKKHLEVMSCPDFIESLVDEKIREF
RCPSILNSAVCDVIDRKCQEHLLKAIINEANRRLFGMKNSSFTMRGNQVLFYTIFSPPKL
PPAASSYVF

CPn_0977 1121329 1122402
No robust homolog present in Genebank/EMBL as of 11/7/98
LYINGFANILKSSFLMEVYSFSEVNTSFOHRVMAALDNMFFLGGRRLKVVSLDSCNSGQ
ACEEYVPISTTEKVLKILSYLLIPIVIIALLIRYLHSNFTAKVSOKFMLKTUGLGIDIK
SFILPGSHVNTMDSATLFKAIRLEGKRVDVEYHRLHSSDKVVFYIPAGKLPDDLRLTHWL
PEKETRKTEYVRHMLAHVMGYLTSGGIKRLDOVX_DSRSSTSLGAEKVLOYRFIDHPOSO
GEFURLLNENITTKGSEDKEVVGSDLFDMAFQGAWPQFTISVIGSPTFSEELYHEMSOKLD
LDGIYPEDDEFEQKFLNTLLKAVLHHGFEGIG/ASMGVIFLICPDSLALQIPFLRNQK

112265+ 1127693

1125/143 CPn_0979 htrA-DO Serine Protease

GIDMITKQLRSWLAVLVGSSLLALPLSGQXVSKKESRVSELPQDVLLKEISGGFSKVATK GIDMITKQLRSMLAVLVGGSULALPLGGGVVKKESHVSELFUDVELKE ISUGFSKVATK
ATPAVVYIESFPKSQAVTHPSPGRRGPYENPFDYFNDEFFNRFFGLPSQREKPQSKEAVR
GTGFLVSPDGYIVTNNHVVEDTGKIHVTLHDQGKYPATVIGLDPKTDLAVIKIKSQNLPY
LSFGNSDHLKVGDWAIAIGNPFGLQATXTVGVISAKGRNQHIADFEDFIGTDAAINGN
SCOPPLINIDGYDYGWTAIVSGSGYYGIGFAIPSLWANRIIDGUIRDGOVTRGFLGVTL
WHILAELAXYYFLEFFFAIXTGVVYYFUFADFASRA, BCLIFANTOREXCOLL MPRINAY
LMCHTRIVLKYSBERFYLEFFTY, AFREC SMAAL, SCHIPFIANTETAFFLAIAE TKGILIISVEPGSVAASSGIAPGQLILAVNRQKVSSIEDLNRTLKDSNNENILLMVSQGD VIRFIALKPEE

1126988 1125504

*similarity to Saccharomyces serevisiae hypothetical 52.9KD

*similarity to Saccharomyces serevisiae hypothetical 52.9KD protein
FYMLINHAKKHAKPYVLIFFSTKDKLSYCDIIFNNCSGKPMNLDSKHFDINSANFLEEFAK
FISFPSISADSDHLQDCENCAHFLVDHVNKIFDVELWETPGHPPIIYASYKSEDPLSPTL
MLYNHYDVQPAQLSDGWXGDPFILREENCNLYARGASDNKOGCFYTLKALGHYYESGONF
FUNIIWLIEGEEESGSIALFTVLEKKKEALRADYLLIVDGGFLSEKHPYVSIGARGIVSM
KISLEEGNKDMHSGVLGGIAYNTNRALSEILSSLHHPDNSIAIEGFYDDLALPSDSDRPD
LPRSDTLRECEENLG/RPGGYEASYSPEESALRFVEINGISGGYTGPGFKTVIPYRATA
YLSCRLVPNQDPDKGAHQVIHHLKGOVPSSLKFSVEILPGGSRGWRSSANLPIVAVLOEI
YSDLYNEECLRLVMPATIPIGPLLGEAAQTSPIICGTSYLSDDIHAAEEHFSMDQLKKGF
LSICQLLDKLPKKE

1127019 1129952

Zinc Metalloprotease (insulinase family)
VTESMKAGDTYRNFIIKSCKDLPEIESKLLEAEHKPTGASIMMIVNNDEENVFNICFRTC VTESKKAGDT/KNFIIKSCKDLPEIESKLLEAEHKETGASIMMIVANDEENVENIGERTC
POTSKVAHVLEHMVLCGSENYPVRDPFFSHTRRSLNTFINAFTGPDFTCYPAASGIPED
FYNLLSVY/DAVFHPLITKGSFLOEAMRYEFNSENHLCYTGVVFNEMKGAMMSGEARLSE
ALNAAIF/SVYTGVNSGGEPREIVTLSHEDVRAFHQSGYSINRCLFYFYGNIKPSENHLDF
LEEKLLBOATKLEKQAVSVPLOKRFKEPVRNILTYPVDHOEDEKVLFGISUMLTCSILEOG
ELLALH/VLEIILMGTDASPLKSRLLKSGFCKOTEMSIENDIREIPMTLVCKGCSPAGAQK
LEALIFASLEEIIREGISENIVEGAVHQLELSRKEITGYSLPYGLSLFFRSGLLKOHGGS
AEDGERINSLFSELRISLKNSDYLAKLIRKYFLDHPHFARVILLPDTELVAKDNKDEQQL
LLSVSEKLTDENKEKIQONVRELTESOEOKEDLNGILENLALDKVPTSGKEFPLIKEGLS
OGS/LHHECFTNDIVFIDVVLDIPPLSGEELPMLRLLVFLMLQLGCGGRSYKEHLEFLLE
HTGGVDVSYDFSPHANNSFLSPSVSIRGKALSSKSEKLCGIVSDMLTSVDFTDIPRIRE
LLMOHNEALTNSVNSPMSYSVSMACSGNSTTGAMSYLTTGLPYVKKIRELTKNFDQNID
FAVVILORLYTKCFSGRRQIVISGSANNYOQLKDNKFYGLLDVILVIFPEMPSINLYV
TSRGLHIPARAAFNALAFPIGDIAYDHPDAAALTVAASILDNVVLHTKIREQGGAYGGA
AANLSRGSFYCYSYRDPEIATTYKTFLKGVSEIASGNFYKEDIYEGALGVVOGLDMFVAP GSRASVAFYRLKSGRIPVLRQAFRRSVLEVTKEHICMVMDKYLESTVQETTLISFAGEEM LRNNVLTLDKDFPIVPAI

1131215 1129962 CPn_0982

CPT_0982
yigh family
KKELASVMNLPVSLACLLLSGCVFFLGVFVSSSLYARKKRAFLEKIQKLEHENQLLQTSL
NLSRHQEQLIEDFSNRLALSSKKLIKDHKEEAQNYFGDTSKSFQSILSPIGTTLTTFKQS
LETFETKHAEDRGRLEGU SQLLAVEKKLEHETHVLTDILKHPSGRGWEGIDLEFILEL
AGMLKYCDYDSGTTSAQGAFRADIIIRLPQDRCLIIDAKAPISDSYFSVEEIDKGDLVDK
IKEHIKTHLSKSYWEKFHQSPEYVILFILPGESLFNDAIRLAPELMEIGASSNVILSSPLT
LLALLKTIAYMMKQENLQKQIQEVSLLCKELHRRLQVVFTHFQKIGKNLNGTVOSYNDMT
SSFQYRVLPTLRKFEGLETSSSHQIEEPTPIESLATSFPHTCDIDTNLAVIESLEKQD

1132045 1131206

CFI_U9993

SIQUETO1-Serine Phosphatidyltransferase
KNPLCYEQKKLWGIDMAGLDLEARGKRRVYTPNAITAFGLCCGLFIIFKSVLRTSSSVEL
FHRLQGLSLLLISAMIADFSDGAIARIMKAESAFGAQFDSLSDAVTFGIAPPLIAIKSLD
GIVYGNFFSSLLLITSIIYSLGGVLRLVRYNLFSQKTVDVSKPYCFIGLPIPAAASIVS
LALFLASDFFPDLPAQLRVGLLSFALLFIGGLMISPWKFPGVKHFRFNVSSFLLVVTIGL AACLFFSGLVDHFVEVFFLVSWLYTLVGFPIFSIIYRKKS

1132370 1135510

CPn_0984 1132370 1135510

"nrdA-Ribonucleoside Reductase, Large Chain"
(KVMYEVEKKYTIVKRNGMFYPFNQDRIFQALEAAFRDTRSLETSSPLPKDLEESIAQI
THKVVKEVLAKISEGQVTVERIQDLVESQLVISGLQDVARDYIVYRDQRKAERGNSSSI
IAIIRRDGGSAKFNPMKISAALEKAFRATLQINGMTPPATLSEINDLTLRIVEDULSLHG
EEAINLEEIQDIVEKQLMVAGYYDVAKNYILYREARARARANKDQDGQEEFVPQEETVYV
OKEDOTTYLLRKTDLEKRFSWACKRFPKTTDSQLLADMAFMALYSGIKEDEVTTACIMAA
RANIEREPDYAFIAAELITSSLYEDETLGCSSQDPNLSEIHKKHFKEYILNGEEYRLNPQL
KDYDLDALSEVLDLSRDQOFSYMGVQNLYDRYFNLHEGRRLETAQIFMMRVSMGLALNEG
EQKNFWAITFYNLLSTFRYTPATPTLFNSGMRHSQLSSCYLSTVKDDLSHIYKVISDNAL
LSKMAGGIGNDMTDVRATGAVIKGTNGKSGQVIPFIKVANDTAIAVNGGKRKGAMCVYL
ENWHLDYEDFLELRKNTGDERRRTHDINTASWIPDLFFKPLEKKGMWTLFSPDDVPGLHE
AYGLEFEKLYEEYERKVESGEIRLYKKVEAEVLMRKMLSML/ETCHPMITFKDPSNIRSN
ODHVGVVRCSNLCTEILLNGSESETAVCNLGSINLVEHIPNDKLDEEKLKETISIAIRIL AYGLEFEKLYEEVERKVESGEIRLYKKVEAEVLWRKMLSML/ETCHPWITFKDPSNIRSN QDHVGVVRCSNLCTEILLNCSESETAVCNLGSINLVEHIPNDKLDEEKLKETISIAIRIL DNVIDLNFYFTPEAKQANLTHRAVGLG/WGFQDVLYELNIG7/ASQEAVEFSDECSEIIAY YAILASSLLAKERGTYASYSGSKWDRGYLPLDTIELLKETRGEHNVLVDTSSKKDWTPVR DTIQKYGMRNSQVMAIAPTATISNIIGVTQSIEPMYKHLF/KSNLSGEFTIPNTYLIKKL KELGLWDAENLDDLKYFDGSLLEIERIPNHLKKLPLTAFEIEPEWIIECTSRRQKWIDMG VSLNLYLAEPDCKKLSNMYLTAWKKGLKTTYYLRSQAATSVEKSFIDINKRGIQPRWMKN KSASTSIVVERKTTPVCSMEEGCESCQ

1135432 1136571

CPI_0985

1135432 1136571

nrdB-Pibonucleoside Reductabe, Small Chain
ISVHKYGGRKKNNPRLENGRERILSITEKRAKMEADILDIKLKRVEVSKKGLVNCNQV
DVNQLVFIKYKMAWEHYLMCCANNMLPTEVIMARDIELMISDELSEDERRYITLINLEFES
TABSILVONNI VLAIFRHITNPEARGYLLMQAFEEAUNTHTELLT.FESDALDEGEVENAYN
ERAGIEAKDDEGMTLTVDVLDEHESVOSSEGLØGFIKNLMISYTIMESITEFYSGEVMILS
FHRONMMTSIGGBYGY LIRDET HILNESIDLINGIKEENDEGMTTHELGEETVALIEKAVE
LETEYAKDLESSMEIDTYPHILADRELERIGLEFITHERNEFTMESTMOLKKEKNEFETRVTEYQTASNIGW

14.967121137395

PÄHEYSMELLÄÄTKMULLA GEMERHIMERS ÄRTEALER TRRUG FORGINANAVONKRIÄNHTAR VARBERFARK MERMININ LOURIG AV LYRELLE LÄKAN KURIL BERGIGEGEL EUNY WELDE PARK MERMININ LOURIG AV LYRELLE LÄKAN ARHI BERGIGEGEL EUNY WELDE PARK ARHI BERGIGEGEL EUNY WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELDE PARK ARHI BERGIGEGER EUN WELLE PARK ARHI BERGIGER EUN WELLE PARK ARHI BERGIN

1137483 1138115 ytgB-like predicted rRNA methylase LENGIFAIGFFMFAYRTLLTHNVVOVSHEIFKTTVVPGDTVIDATCGNGNDSLFLARLLQ GEGRLVVYDIOKEALSNALLLFETHLSEQERSVIEMKEQSHEHILEKDVKLIHYNLGYLP KGNKEITTLARTTEISLEYALNIVRPDDLITVVCYPGHPEGEKETHSVESLAQRLHPKEW CVSSFYVANRCRAPRLFIFOROGSESSVDKG

| 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 VI FILIS INVPFLI ICKGSNCLFDDRGFOGFVLYNA IYCKOFLEDAR IKAYSCLSFALG KATAYNCYSCLEFAAG I POSVGGA I FINIAGTHESDI SSVVRIVET INSEGELC SYSVEEL ELSYRSSRFHRQOEFILSATFOLSKKOYSADRSKS I LOHRLHTOPYTOPSAGC I FINIPEG TSAGKLIDAAGLKGLA IGGAQ I SPLHANFI INTGKATSDEVKQLIA I I I QSTLKTOG I DLE HEIRIIPYQPKIHSPVSEK

1139552 1139016 CPn 0989 CFBJ2 hypothetical protein LRTSLAVKCVLLTIFWILLVMATLSPEKFSGSPISISKEFPQQKMREIILQMLYALDMAPS AEDSLVPLLMSQTAVSOKHYLVALNGYKSILEKSQELDLIIGNALKNKSFDSLDLVEKNV LRLTLFEHFYSPPINKAILIAEAIRLVKKFSYSEACPFIQAILNDIFTDSSLNENSLSI

CPn_0990 1139880 1140440
infC-Initiation Factor 3
svalnrkinngiraprvrligsageolgilaikdaldlareagldlvevasnseppvcki
mbygkyryglitkkekdskkaohovrikevklkpnidendfstklkoartfvekgnkvkit
cmrgrelaypehgfkvvokmsogledigfveaepklagrslicvvapgtvktkkkoeks

CPn_0991 1140394 1140612 rl35-L35 Ribosomal Protein KQRKNRKSLMPKMKTNKSVSARFKLTASGQLKRTRPGKRHKLSKKSSQEKRNLSKQPLVD

KGQVGMYKRMMLV

CPn_0992 1140622 1140996 r120-L20 Ribosomal Protein GKLVMVRATGSVASRRRRRILKQAKGFWGDRKGHIRQSRSSVMRAMAFNYMHRKDRKGD FRSLWIARLNVASRIHSLSYSRLINGLKCANISLNRKWLSEIAIHNPEGFAEIANQAKKA

CPn=0993 1140975 1142030 phdS-Phenylalanyl tRNA Synthetase, Alpha KSPSSHSLGIRISMEMKEEIEAVKQQFHSELDQVNSSQALADLKVRYLGKKGIFRSFSEK LKGETDKAKLGSLINDFKTYVEDLLQEKSLVLLASEQAEAFSKEKIDSSLPGDSQPSGR HILKSILDDVVDIFVHLGFCVREAPNIESEANNFTLLNFTEDHPARQMHDTFYLNATTVL RTHTSNVQARELKKQQPPIKVVAPGLCFRNEDISARSHVLFHQVEAFYVDHNVTFSDLTA ILBAFYHSFFORKTELRFRHSYFFFVERGIEVDVSCECCGKGCALCKHTGWLEVAGAGMI HPOVLRNGNVDPEIYSGYAVGMGIERLAMLKYGVSDIRLFSENDLRFLQOFS

HPOTERION OPPETYS GYAVGMGIERLAMLKYGVSDIRLFSENDLRFLQGFS

TOPT 1994 1142371 1144440

CTB37 hypothetical protein
LFWERINGRIMKRSRRNFEQALENLEKLKEISLATSNDSYLINDPARFNORKOTGSSVMENK
PALIAVENYLLEISCVSKSHADKALKESDFLIAGVONVFSFLENQEDLYKSLLDEYSEVT
KAYPBEVKKNILKEVPTYDLSTDEETEEHKEPECFLINILVEVKRDRSYELFYMLDEQDKFY
NDÄLÄGI IYKONKLHETVINGSDPLITTLLIMNSEEVENIASSLUVINDMPLIELFYGYALSH
LDIEAVVKVHNAVMALFFSRYEATMVFKSPKKHNIWYFNDFLLFLREAWKDLINDWIDSQ
ERKOTKLLASALSIGIFESKLVFEEASRYLYFNIOTKLENANGKKPLSPGQYLYDAVEEL
HRLESKYPNOFLFKANDRVLEHESRYDPMHIGILDFSLEGTLKLIGKSIDII JKSPSPVTQ
SSILYANCNEEFLGFLNAKAHRSEVTLVLNIQNRISRKERARSRVIEEALECEHAPYVH
AFSEPEPEELLGULESIHGDIETFADFFSILQEEFHKEPLLASSFFLTKELKEFVGSSILKE
KLTALKOHFFAKKKILFRNDKLLLLHLLSYLIVFKLIERTNPRISTVVSKOGLDYVSVFI
AGFÄFFSREAFWDEHSLKLLLTNVLSPTLVARDRLVFVSHIELLSKFVNCLKKNRQGFSS
LKSFFKNDIEGWEFTGYLHELTEVSHKHNL

Philopos 1145515 1144415

CPT-10995 1145515 1144415

CPT-10995 1145515 1144415

CPT-10995 1145515 1144415

CPT-10995 1145515 1144415

RMLEWKRHLUTREWFALTSLUVLALIFYASIHHSLHTLKGASTAASGASVKLSILYYLAG

ISÉKÁEFLMPQLVAVATTSTLFAMQNKREIILLQASGLSLKSLMHPLLLSGAVIMMYLYA

NFOWLHPICEKISITKENMDRGTTDKEGGKIPALYLKDOTVILYSSIEPKTLTLNNYFWI

KDPKTIYTMEKLAFTTLSLPIGLNYTOFFANDSENLELKEFFDWKEFPEIFFNFYENPFS

KLFSAGNKNRLSEFFKAIPWNATGLGLSTOVPQRILSLLAGFYYVLISPLACMAAIILSA

YLCLRFSRTPTVTLAYLIPLGTVNIFFVFLKAGIVLASSSVLPTLPVMAFPLIVLFLLTN

YAYAKLQ

CPn_0996 1146592 1145519
CT839 hypothetical protein
AMPILMKVLIFRYLKTAAPCTLSLICISIISSLQE/VAYIAKDVPYDTVLRLMAYQIPYL
LPFILPGSCFVSAFSLFRLSDNNHMTFLRASGAGOSIIMFPVLMVSGAICCLMFYTCSE
LASICRYOTCKEIANNAMTSPALLLOTLOKKENNRIFIAVDHCAKSKFDNVIVALKGNNE
ISHVGIIKŠIIPDTTKDTVKAKDVVFISKLPD9LTESSSPSSQRFYIETLDELLIPHITS
TLFAGKSYLKTRTDYLPMKOLVKQSLKHSHLBETLRRVAIGFLCITLTYAGMIGIHKPR
FRKSIALYFIFPILDLILLIVGKNTKNLPLAFMLFVFPQLVSWVVFAARAYRESRGYA

CPn 0997 1146699 1147664 CPD_0907 1146699 1147664
mesJ-PP-loop supertamily ATPASE
AYKWULSDLLRDKQLDLFFASLDVKKRYLLALSGGSDSLFLFYLLKERGVSFTAVHID
HGWRSTCAQEAKELEELCAREGVPFVLKTLTAEECGDKDLENQARKKRYAFLYESYRQLD
AGGIFLAHHANDOAETVLKRLLESAHLTNLKAMAERS*VEDVLLRPLLHIPKSSLKEAL
DARGIGYLOPPINEDERYLRARRKKYLFPWLEEVFCKHITTPLLTLGEESAELGEYLEKQ
AQPFFGAATHQDCGELPCPDCLIQOAFLCKWMKKFFNNAGIAVGRHFLOMVYDHLDRS
GCATLEMRNKIVIIKDJVVVID

1147 31 1150584 GET_0.008 1147/31 1150584

LIGH-ATP-dependent zinc diotaise

LIGH-ATP-dependent zinc diotaise

LIGH-ATP-dependent zinc diotaise

LYNIRE IVPEDSHKIALDOM-FCCREROVOTQEGGLAFHYLELIDOGHREDEDLQETS

KSHITTI/SKEVING HEMPSYLGGSPIPEGGYA ESYPCE/CGSVHITEDLATEARTPOLINL

BEDJERYPTEGRIPPALBYYGEDLYELICKYESPVI/ITGSFTLKREENDLYGQVEVSLTQ

ETUPFAAYTLYGOVI-SPUNK ESSCLVVSEOGERFSQLROVRLYREEMNKYHKLVFARDLN

QAQLEKLRGELSOTVWYFNNQELSSRSLEKQDPEVF/HMFAGAKEEMTAFKFNHGLGFKA 47
PULIPHINLULEKTEKSGEPS
PARMLLKOONKVTFADVAGI
ELIEIVDFLKNENNETSLEGRIPKOVLLIGPPOTC
KTLIAKAVSGEADRPFFSIAGSDFVEMFVQVDASRIRDMFGOAKRAPGIIFIDEIDAVC
RHRGAGIGGGHDEREOTLNGLLVEMOFOTNEGVILMATNREDVLDKALLRFGRFDRRV
VMNLPDIKGRFEILMVHAKRIKLDETVDLMAVAJSTFGASGADLENLLNEAALLAARKDR
TAVTAVDVAEAADKVLYGKERRSLEMDAEERKTTAVHESGHAVVGLCVQHGDFVDKVTII
PRGLSLGATHFLPEKNKLSYWKKELYDOLAVIMGGRAEEIFLGDISSGAQQDISQATKL
VRSMYCEWGMSPQLGNVTYDERSOCLTGYGGYHEKSYSEETAKTIDTELRMLLDAAYQRA
LBILMBRABLELMTMAIREDIA-BROVKEIMDHTADFERRRAHLPBEIMLFKKSSCEL
PREPERBOULGGAREIAT

CPn_0999 1152859 150766
pnp-Polyribonucleotide Nucleotidyltransferase
QETFMNFQTISINLTEGKILVFETGKIARQANGAVLVRSGETCVFASACAVDLDDKVDFL
PLRVDYQEKFSSTCKTLGGFIKREØRPSEKEILVSRLIDRSLRPSFPYRLMQDVQVLSYV
WSYDQOVLDPDLAICASAALAIBIPPGSNIVAGVRIGGIDNOWINPTHTELASSTLDL
VLAGTENAILMIEGHCDFFTEEQVLDAIEFGHKHIVTICKRLQLWQEEVCKSKNLSAVYP
LPAEVLTAVKECAQDKFTELFYKDKKVHAATAHEIEENILEKLQREDDLFSSFNIKAA
CKTLKSDTMRALIRDREIRAGGSLTTVRPITIETSYLPRTHGSCLFTRGETGTLAVCTL
GSEAMAGRYEDLNGGLSKFYLQYFFPPFSVGEVGRIGSFGRREIGHGKLAEKALSHALP
DSATFPYTIRIESNITESNGSSSMASVCGGCLALMDAGVPISSPIAGIAMGLILDDQGAI
ILSDISGLEDHLGDMDFKYAGSGKGITAFQMDIKVEGITPAIMKKALSQAKQCCNDILNI
NNEALSAPKADLSQYAPRIETMQIKPTKIASVIGGGGKQIRQIIEETGVQIDVNDLGVVS
ISASSASAINKAKEILEGLUVGEVEGKTYPGREVTSVVAFGAFVEVLPGKEGLCHISECSR
QRIENISDVVKEGDIPDVKLLSINEKGQLKLSHKATLE QRIENISDVVKEGDIADVKLLSINEKGQLKLSHKATLE

1153193 1152891 CFI_TOOV
TS15-S15 Ribosomal Protein
SAFAAIILRRHMSLDKGTKEEITKKFQLHEKDTGSADVQIAILTEHIAELKEHLKRSPK
DQNSRLALLKLVGQRRKLLEYLNSTDTERYKNLITRLNLRK

1153369 1153869 yfhc-cytosine deaminase
yylelcgeklinmekdiffmooafkearkayDodevPvgcvivkDDkiiarahnsveklk
DatamaeilcigsaaODLDnwrllDtvlyctLepclmcagaiOLariprivwaapDvrlg
aggwwnifteeHpfHtvsctggvcseeaeHlmkkffvekrreksek

rn 1002 1153844 1154089 TT845 hypothetical protein
KSAERKVKNKIVTLLDQLYEDQESRLQKLGEEIVPNLTPEDLLQPMDFPQLEGNPAFRFE
EGVLSGIGEVRAAILAALSQEN

1154862 1154092 CT846 hypothetical protein
T\$NKTIHPLLWGPDRQIAGKASMRVIFPDKHNNFPNLSKLLKKLPSVILVTSCIAPFFSY
HINKFFGIPGLLEILALSVKGIQKHHFWOFLTYPLITADSLSLNKDQSFEITQRLLLRNV
LDFFLFYKAIQHLIRKIGAFSVLVVISQQALIIGAVLWGFMALIKSSQSFFGPESIIGGV LTVQIFLDPEKRFTIGPTPLSVSIKWGFLFVLGFYCCILIFSGAFLLLLASMLAIVLAIL FCKKEKIPNPYTTSLRF

1155418 1154879 CFN_1004
1153416
1154679
CT847 hypothetical protein
HLSIEELMSIQPVSNTTTKADKVIPDSTKVISDSITINKQSAFYFCISVMLRLSESTTEY
GKSILAVLEDNTIVQQQRVKELINLPLLKVPDLQKKDGSDDEYKNQNEIQAYQSSNQQIS ANROMIQQELSSAQQRAQANQKSVNSTTIESMQILQATSSMLSTLKELTIKANLTNSPSD

1155957 1155415 CFR_10US
CT848 hypothetical protein
NRKPVRLMMIIIDPLSAKKPLQAAINVPGTPITGPNTATADDIIAKFSKDSNPLIVTVY
VYVOSULVAQDNLSIIAQELQANSSACTYLLNNQSALYQVVSIPKNKLMDNSSSYLQNIQS
DNQAIGASRQAIQNQISSLGNAAQVISSNLMTNNNIIQOSLQVGQALIQTFSQIVSLIAN

CPn_1006 1156493 1155990 CT849 hypothetical protein TKVNFFIMSITTLGTLPTVNTINSSRPPLEPLNTPKIGAVLFSIYELLLQAIEIRQQTVL TQSQQLNDNTNIQQQLNQETNQIKYAIVSAGAKEDEITRVQNQNQNYSAQRSNIQDELVT TRQNGQIILSHASTNINIIQQQSSQDSSFIKTTNSIGSTVNQLNKPLG

1156689 1156907 CT849.1 hypothetical protein
LWYKSLAGEEKDVSGNECNDYPEVFKDDVSAYVLVTCGQMSSEGKIQVEMTYEGDPAVIS
YLLTKARDSLDES

1156904 1158223 CFM_1008 115904 115823 CTB4850 hypothetical protein VLNYSFIGMLKPMYVLSKRLYRWVNQLIKLGDLVKNSRSFSVEWVFISALLLIFGCLGCA SVVKVSLVPFLLLFSFLAFFLILCFRGKGYALLLGVFYTLYVAKYVVGETLYVSFWLSGL GVSFLLAFGLFLGCWULAGEEMVKREGPLRISEDLDAGRGSYEDLLTKSGEKEFLDAR AQGLDRELTECQELLKAAYQKQEYLTIDLKILADQKNGWLEDYAELHNKYIELVSKNGDV VFWVAEPSVGESOSERVDVSRWYSALQEKEESLERLRNETLVEKQRCSDYEHRCQELG LLLQNFTALERRCEELQNLLNQKETQINELHQLVCKSEEKVSVEPSAHAETSCVEEKQYK GLYSQLQEQFLEKSETLSLVRKKLFAVQEKYLTLKKKEELTKQDISFDDISMIQGLLERI EILEEFYSHLEELVSRSLSL

CPn_1009 1159085 1158186 map-Methionine Aminopeptidase YRLLHRYTLMKKNDPCWCGSGRKWKQCHYPQPPKMSPEALKQHYASQYNTLLLKTPEQKAK TYMACQITARILDELKASQKOVTINELDELSQELHKYYDA LAPPHYSSPPPRKTICTS
LIEVTCIPISIENDI PLKUGINNI IDVISCIVOGYYGCOSPWM IGEVPE IKKKICOALECL
NDSIGALIKO IPLIKUGINNI IDVISCIVOGYYGCOSPWM IGEVPE IKKKICOALECL
NDSIGALIKO IPLIKISEA LERADOTYPOVODOVYKIYO I EPIENPYVPHYRNSMI P
LAICMI IPTI EPMINAKKESVVDPKNOMEARYCONOPISAOMEINTA I TETTYE I LITLILD

CHR_1019 LISSEYS 1159067
CTROOF TOPOTHER FOR THE PROTECTION
WHILLIALD CLEEP VERY REPRESENTATION OF THE CLEEN ALVAL LEPVING
CHROCELLARY OF THE CHROCE SHAMLAR IMPRESENTATION OF THE PELAPPY TYPOTA
VITALLEY MERGITY ORDER CHROMITTAWAR OLD CHECK CONTROLLAR OF THE CONT DEMOVMEMBERGED LAFNIGEVICE

Chu_1611 1160309 1159902 1154402

119

CTRS : hypothetical protein EMRLKNYPMIQESFELPOTCILLLASDSLTNILALHH AMFALYGLAL/GLKVLNTPVCATEVVCGIAVTLAGVRA VKORML/LLRESFFAFI CKEESWI PYKFNMS PSYS PC LOP LALPLMFGPSG

CPH_1012 1162220 1160421
yz+B-ABC cromsporter permease
ALFOLITSYMKKKFIFYFYTVFSLLFLWEMTSRHPPTFSFFCPPPSGIASSTLQSLPLLL
rowmenuskalsssericatenthologistics. State 14 december 14 december 15 december 16 december 17 december 16 december 17 december 16 december 17 d HIFSGLKIAIGSAGFAAIAGEWVASOSGLGILMLESRRNYEMELAFAGLATLSILTLSLF QITLLIEKLIFSLFRVKRMSLKHKSVAKKALSVLALIPIMLIPWKGNSKSPPDKKNLTSL TITLDLIERTIFSERVKRSSLARRSVAKALSVLALIFHLIFMANASEPURNITSL TILLDMTPRINITPLYAGVAKOYFKOHGIDLQLQKNTDSSSAVPHVLFEQVDMALYHALG IMKTSIKGHPIQIVGRLIDSSLQGFLYRSQDPIYKFEDLMGKVLGFCLMNSRDLMRLLET LNRNGVVPSEVKNVSSDLISPMLLNKIDFLYGAFYNIEGVKLQTLGMFVKCFLSDTCDLP TCPQLIVTFKKGTKASEPEIVEAFQKALQESIIFSKDHPEDAFKLYAKETKSIPKNLYQE YLQWEETFPLLAQSQDPLSKDLVDKLLETIIKRYPELASEVAKFSLNDLYNPSLPEEQSV

1162209 1163624

CPH_1013 1162209 1163624

fumC-Fumarate Hydratase

RENSLAHRGNIDMRQEKDSLGIVEVPEDKLYGAQTMRSRNFFSWGPELMPYEVIRALVWI

KKCAAQANQDLGFLDSKHCDMIVAAADEILBEGFEEHFPLKVWQTGSGTQSNMWNDEVIA

NLAIRHHGGVLGSKOPLHENDHVNNSQSSNDVPFTAMMIAAVISLKNKLIPALDHMIRVL

DAKVEEFRHDVKIGRTHLMDAVPMTLGQEFSGYSSQLRHCLESIAFSLAHLYELAIGATA

VGTGLNVPEGFVEKIIHYLRKETDEPFIPASNYFSALSCHDALVDAHGSLATLACALTKI

ATDLSFLGSOPRGGLGELFFPENEGSSIMPGKVNPTOCEALGWCAQVLGNNQTVIIGG

SRGNFELNVMKPVIIYNFLQSVDLLSEGMRAPSEFFVKGLKVNKARLQDNINNSLMLVTA

LAPVLGYDKCSKAALKAFHESISLKEACLALGYLSEKEFDRLVVPENMVGNH

CPn_1014 116545 ychM-Sulfate Transporter 1165456 1163732

ychm-sulfate Transporter
ALASTLGYCIVXVPMAFKNFIPKLYTSIKEGYSFNTFKKDFQAGITVGILAFPFAIAIAI
GVGVSPIGGLLASIIGGLLASAMGGSNVLISGPSSAFISILYCLSAKYGAEALFTVTLLA
GVFLIAFGLTGLGTFIKYMPYPVVTGLTTGLAIIIFSSQIKDFLGLQMGANIPADFLPKW
IAYWDHLMYWDSKSFAVGLFFLLINIYFRNYRFRYGWNIAIVTATTLWHLKIDIPTIG
SRYGTLPTAIPLPKIPOLSTKILGLMPDALTIAVLSGLETLLSAVVADAUTGWRHQSNC
QLVAQGVANIGTSLFSGIPVTGSLSRTAASIKSEATTPIAGIVHSIFICFILLLLAPLTV
KIPLTCLAAVLILIAWNNSEIHHFIHLFTAPKKDIVVLLTVFILTVMTTITAAVQVGMML
AAFLFMKQNSLSDVISTAKYFDKDSDFLSKAEVPONTEIYPINGFFFFGIADRLKNLLN
DIEKPPKIFILCMTRVPTIDASAMHALEEFFLECDRQGTLLLLAGVKKTPLADLKRYHLD
ELIGVDHIFSNIKSALLFAQALTNLESKTSTRHLV

CHILIO15 1165550 1166893
CTG\$7 hypothetical protein (possible IM protein)
KNNEGNFSFFTSVRVRSKVDHEIILEVTHLKLQLCALFLFGYLAIVFEHIVRVNKSAIAL
ANGGLMWLVCFSHIPMADHMILVEEIADMSQVIFFLFSAMAIVELIDAHKGFSVIVKFCR
IQHTLLLWALIGLSFFLSAALDNLTSIIIISILKRLVKAREDRLLGAICVIAVNAGG
AMMHGDVTTHMWINNKITSWGIIRALFVPSLVCVLVAGFCGOFFLRKGSTLIAKDVE
LOSAPPKSLMIFFIGISSLLMYPVKAKCLGLPFPFMALLGGLUVALTSDWIHSPHGEDRY
HL#RYPHILTKIDISSITFFIGILLAVNALSFANLLTDFSLWMDKIFSRNVVAIVIGLLSS
VLDNVPLVAATMGMYTLPLDDTLWKLIAYAAGTGGSILIIGSAAGVAFMGLEKVDFLWYF
KREMWILLASYFGGLFSYVLZSLNFFI Krīzwialasyfgglfsyfvleslnffi

1167027 1168898

CPETO16

1167027 1168898

CT\$50 hypothetical protein

KREMFMKKGKLGAIVFGLLETSSVAGFSKDLTKDNAYQDLNVIEHLISLKYAPLPWKELL

KREMFMKKGKLGAIVFGLLETSSVAGFSKDLTKDNAYQDLNVIEHLISLKYAPLPWKELL

KREMFMVONCOTQQARLQLVLEEKEPTINNCQKVLSNYVRSLNDYHAGITFYRTESAYIPYVLK

LSEDCHVEVVDVOTSQQDIYLDGEILEVDDMGIREAIESLRFGRGSATDYSAAVRSLTSR

SAARCDAVPSGIAMLKLRRPSGLIRSTPVRWRYTPEHLIGDFSLVAPLIPEHKPOLPTQSC

VLFFSGVNSQSSSSSLFSSYMPYTWEELKVONKQRFDSNHHIGSRNGFLFTFGPILWEQ

DKGEYRSYJYKAMDSQGNPHRIGFLRISSYVWTDLEGLEEDHKDSPWELFGEIDHLEKE

TDĀZIIDOTHNPGGSVFYLYSLLSMLTDHPLDTPKHRMIFTTQDEVSSALHWQDLLEDVPT

DEDAMAVLGETMEGYCMDMHAVASLQNFSQSVLSSWVSGDINLSKPMPLLGFAQVRPHPK

HQYTKPLFMLIDEDDFSCGDLAPAILKDNGRATLIGKPTAGAGGFVFQVTFPNRSGIKGL

SLĀSSLAVRKDEFIENLGVAPHIDLGFTSRDLQTSRFTDYVEAVKTIVLTSLSENAKKS

EEGTSPQETPEVIRNSYPYTTSAS EEQTSPQETPEVIRVSYPTTTSAS

VPADLINNPEDIDTNIVNHSGDIAMTAGASTPEDVVQACIRKLSSLIPGLQVEND DVVFQLPKELRCS

CPn_1018 1169895 1170629
No robust homolog present in Genebank/EMBL as of 11/7/98
RMSYFN7QKNSVVLRSLGLLAKFFSRLLYRVFFSFREGIYLFSSLYLKYPRFFFYDLGKY
VYSLRHCPYAKLGRLPGASLLKEGNVYGETPWSVLAKICQAFDITSQDILYDLGCGLGKV
CFWFSHVVRCQVIGIONQPHFIRFSSNMHRKLSSGFALPDTEEFKNVVLSASYVYFYGS
SFSRRLLNEIILKLSEMAPGSVVISISFPLDSFSRGKECFFTEKSCSVRFPWGKTIAYKN
1BKC3

1172146 1170638

CFn_1019 1172146 1170638
CTH90 hypothetical protein
LHRPNIMTVSYQSISTPPPEGEPDIFVDCNATEEAV/AAEVQVALENGEQYAMLRATSEL
CFG LTTGGECALTQALEPREKELQEEQFLVKNGILMESTSLENLKFGQSQCTSLASHRNP
LAQQCTSCNSTTKALTETTSSSEPFESCAPECDSVJVKTETVSYOTPKAQEQEASAASQ
GOADHIVRSVSGGT IKEBGAKEKVSQSTKGAETQKHTQTKSDATLSPMSLYSTLIKEVPO
ALSSTKSQQKBEEHRBURQOEQYEQEQEQEERKKNTPMCTVESJAVTSSNGY/ESYTPI
LPDF1VFFALGSQLSVLAGKRVTNLDVLRICTELMFLMLKGVADTMTREEERELMERE
AHELAASTSRQAKYAKWISIATATLGILSIA JAPNYGE ISGDZYLSFVORISGPFKDATAK
TEFEKGI GVWTTCLSOLOFEAASKWUELSFSAVEAUS GEBERGMERE TELKO DIKALLOPOLEVYORKAHETRESAARAARAKEAKKAA KWODEALLA LEEAKDUM KOMUNETTINT FÖLEHPVARGITAÖ

117 604 - 1172150

CTM_1020 117 0-4 1172150
CTM_1 hypothetical professioners by devaloraçkolpkeftok
Egockober fuhktet laterer lergosfessork byd devargeackolpkeftok
Egockober fuhktet laterer lergosfessork byd devargeackolpkeftok
Egockober fuhktet laterer lergosfessork
Egockober fuhktet laterer lergosfessork
Egockober fuhktet laterer lergosfessork
Egockober fuhktet laterer lergosfessork
Egockober fuhktet laterer laterer lergosfessork
Egockober fuhktet laterer laterer lergosfessork
Egockober fuhktet laterer lat

DSECOELLOSRREERSETY CHEKKIETYVQIKDLUKDLFSQDQDSNCHQAK DTGRKNRIAKAAQAVPVIPA VFTLSYULTKOGILJSDFISYOCHKOSVESTORELDA LHEKRIETIKVSIEKEKRERLUGSLSDIIGWLAPFVSIGIGIVAILSCOGIFAFAGFFAG LISLVIKCLEKLKFVOWLEKHLPIKNEELJARIITIIGWVVLTFVILSICTLKVENLGF SPIIEGAIRGIPAIESTMAALRCAILESOAEIYKLKGKLTKIQLDIELKSFDRDDHYER SQELLDNMESSFEALSRILMYMRELDQV/LHSLRG

CPn_1021 1174270 1173698

Peril flow (Nail Leas Roll forme Thomas Home Hillary Formas Hillary PK DALICKFOFOFOFOPOLEGYLEGYZETKNIALO TYCK IFKISSEDLEKYYKEUYHAYLOKDYAKU ITVFRWLVFFNPFVSKFWFSLGASLHMS EQYSQALHAYGVTAVLRDKDPYPHYYAY ICYTLITNEHEEAEKALEMAWVRAGHKPLYNEL KEELLDIRKHK

1175/09 1174216

CPD_1022 1175/09 1174216
CT863 hypothetical protein
FSFFFYALKLQIMMMPVPSAVPSANITLKEDSSTVSTASGILKTATGEVLVSCTALEGSS
STDALISLALGGILATQGE/LLQSTNVMQLLFLPFEVVELE:QUVDLLVQLEHAETITS
EPQETGTOSRSEGTLPQQSSKSQSALSPRSLKPEISDSKQQALGTEKDSAVRKHSEAPS
PETQARASLSQASSSSGSLPPQESAPERTLLEQQKASSFSPLSGFSAEKQKEALTTSKS
HELYKERQDRQQREQHDYKHDQEEDAESKKKKKKRGGLGVEAVAEEPGENLDIAALIFSD
OMRPPAEETSKKETTFKKLESPMSVFSFFIESKNPLSVGSSIHGPIOTEVLDNVTERH
KLMARILGQAEAEANELYMRVKQRTDDVDTLTVLISKINNEKKDIDWSENEEMKALLNRA
KEIGVTIDKEKYTMTEEKKLLKENVQMRKENMEKITQMERTDMQRHLQEISQCHQARSN
VLKLLKELMDTFINNIRP VLKLLKELMDTFIYN\(\mathbb{R}\)P

CPn_1023 / 1176008 1176331
No robust homolog present in Genebank/EMBL as of 11/7/98
GLDFLEIFIMKKYVTLSIIFFATYCASELSAVTVVAVPLSEAPGKIQVRPVVGLQFQEEQ GSVPYSFYYPYD GYYYPETYGYTKNTGQESRECYTRFEDGT I FYECD

CPn 1024 1177317 1176334

CPn_1024 1177317 1176334

xerD-Integrase/recombinase
IFFFPWFSL/SLKIAPLPILKLHSLASMTMPSTQFHTTILEQFSLFLSVDRGLCQQSIAA
YRQDISSFL/TISAISSPQDISQNSVYIFAEELYRRKEAETTLARRLIALKVFFLFLKDQQ
LLPYPPIIEHPKIWKRLPSVLTPQEVDALLAVPLQMENPRHLAFRDTAILHTLYSTGVR
VSELCDL/BLGHVSDDC IRVTCKGSKTRLVPLGSRAREAIDAYLCPFRDQYQKKNPHEDHL
FLSTRGHKLERSCVWRRIHNYAKQVTSKPVSPHSLRHAFATHLLDNKADLRVIQEMLGHA
RIASTE/YTHVAADSLIEKFLAHHPRNL

CPn_1025 1177266 1178879
pgi-Glucose-6-P Isomerase
GAEQFSSYREKTMERKRFIDCDSTKILQELALNPLDLTAPGVLSAERIKKFSLLGGGFTF SFATERLIDAILAALISLAEERGLHESMLAMOGGVVNYIEGFPSEMPALHTATRAWT DSFTTEAEDIAVRSRVEAORLKOFLTKVRSOFTTIVOIGIGGSELGFKALYRALRAYCP TÜRKVHFISNIOPDROASEVLOTICOAKALVVVVVSKSCTTIETAVNEAFFADYFAKKGLSF VÖHFIAVTCEGSPMDDTGKYLEVFHLWESIGGRFSSTSMVGGVVLGFAYGFEVFLOLLO SANDO IALOPNARENLPHLSALISIWNNNFLGYPTEAVIPYSSCLEFFPAHLOOCCHES JOKSEAOOGRAVGFSTSPVIWGEPGTNGOHSFFOCLHOOTDIIPVEFIGFEKSOKGEDIS FOGTISSOKLFANNIAOAIALAGSENTHPNKNFDOMPSSVLVSSOLMPYSIGELLSYY ENKLVFOGLOWGINSFOQEGVSLGKALANRVLELLEGADASNFPEAASLLTLFNIKFR

n 1026 1178961 1179137

CSFGFGKICEDRMFFIAVRSRGFLDIHGILAARKGKQVVKSTAGAWIGSRGAVFYSLVS

1179172 1180755 No robust homolog present in Genebank/EMBL as of 11/7/98 NMPGSVSSPPLSPVIVRERVPSSSGSDLIQPHAVLKISILIFALVTILGIVLVVLSSALG

MMGSVSSPELSPVIVRERVPSSSGSDLIQPHAVLKISILIFALVTIGIVLVVLSSALG ALPSILVITVSGCIAIAVAGLIGIGIULTVRLILESTIRKVDAMSVDAAVKEGVJLSKIRELSE ENREIRDRNRAVEDQCAHLSEENKDLRDPEYLHGMTERLIASLEIENQALVAENILLKDW NASLSRDFRAYKQKFPLGALEPWKEDIACIMEQNLFLKPECIAMVKSLPLETQRLFLYPK GFQSLVNRFAPRSRFFOTPKYEYNSRIENEDGKVAAVCARLKKEFFSAVLGACSYEELGG ICERAVALKETLPLEAVYDTLVQEFPNLLTAESLWKEWCFYSYPYLRPYLSVDYCKRLF VQLFEELCLKTTGSPEDQALVRLFSYYRNHIPAVLASFGLPPPTGGSVFVLLPKQEN LLWSQIEVLATRYLKDTFVRNSEWTGSFEMMFSYNEMCKEISEGRIFFAEDYETRHSEEF PPSPLSEEGEGEFLPPCSEEVSVLERPDLDVDSMWVWHPPVPKGPL

1180995 1181999

mdhC-Malate Dehyrogenase FFLKGVRMAFKEVVRVAVTGGKGQIAYNFLFALAHGDVFGVDRGVDLRIYDVPGTERALS GYNMELDGGYPLLHRLRVTTSLINDAFDGIDAAFLIGAVPRG GWERGDLLKONGOIFSL GOVALNTAAKROAKIFVVGNPVNTNCWIAMKHAPRLHRKNFHAMLRLDONRMHSMLAHRA EVPLEEVSRVVINGNHSAKOVEDFTOARISGKPAAEVIGDROWLENILVHSVQNRGSAVI EARGKSSAASASALAEARSIFCPKSDEWFSSGVSGHNPYGIPEDLIFGFPCRMLPSG DYEIIPGLPWEPFIRNKIQISLDEIAQEKASVSSL

CPn_1029 1181987 1182844

No robust homolog present in Genebank/EMBL as of 11.7/98

RVFYISTMLWGVSHROSFDELSQNAFKNIFNKQRFCFIFCSLCCFGFVFALFLKLCSRLA

PEISLSTLGLGAFFCAFSVICASAIIVOFLLHKESGOETISKLCCAIKNTWSSLWLSLLVS

MPFFIANVAVVTVAMLSSFLGSLPWVGKLFHTVLIFIPYLSATALILLELSFSCLFFCI

PVLHNQESIDYRKLLECFRGNILROFIGVVIALVPLALCSWLALDSFYLMTHLVEIADIH

TWSFLAOMFVLIVPIALILTPAVSFFFNFSFSFYLAKQEEEKALVK

1183901 1182843

CPH_1030 1183901 1182843
predicted D-amino acid dehyrogenase
FKVHFMRIAVGAGYAGLSVTWHLLLHSOGTATIOLFDPIPLGEGASGMSSGLLHAFTGK
KALKPPLADQGINNTHALITEASKALNVPIVISGGILRPAIDEDQAQLFTERVEEFPKEV
EMWEKARGEISIPSMVIPRAUGALFIKSOVTLHRIGLY (COLADACHKLTTQFYDELIEDL
ADIEEFYCHIIVTTOANASILPELKDMFVNKVKYOLLEIJAMPKLAHLSFSINAHKYMV
NTQKNTCILAGATFEHNQPEETPDPAIAYQEIMPPVLOLFFGLKDAQVLHCYALMRSSKS
RLPVISRIREKLMFIKSLOSKGLLYHGITTOMLAGAVLRKSTAYIAKEFLFTI

CPn 1931 1185567 1184098

CENTRAL TATE THE STATE STA

CLITAVISCULSWITTIVAETPEGAAKNOTEPETETTE VSLYITSSVMOLAML VYFOSNAWITMLS ITTYMVLPAYLASAAFLFKLSKSK SLWL IYAUGLKYLFMALVLLALG IPFYIDAGKKKKNAK TKAPLAMITGILGVVY

CPn_1032 1186153 1185566 CT173 hypothetical protein LMAY/TPYPTIAFHTOGIGEODOMPPOPFETFCYDSALLQAKIENFNIVPYTSVLPKEL POTE OF THE POOR STANDARD FOR A PARTICLE OF THAT A PART OF THE PARTY O

ALGELNEENAEPAKVN

1187656 1186187

CPD_1033 1187556 1186187
CT372 hypothetical protein
NNKKKDYSGEFLTTDTVDSIAFLPSEENFCYIKTILFFRVKKKHYAFFYGEFMISFRFLL
LSGLCALGISSYAETPKETTGHYHRYKARIQKKHPESIKESAPSETPHHNSLLSPVTNIF
CSHWKKGISVSNLLTSVEKATNTOISLDFSILPGMFFYPHKALGGTQALEIPSWGYFFSP
STTMTLVSBPTAGGIVDFSYTLIHYWQTNGVDANQAAGTASSMDVSNRENNLAQLIFS
OTFPGDFLTLAIGQYSLYAIDGTLYDNDQYSGFISYALSQNASATYSLGSTGAYLQFTPN
SEIKVQLGFQDSYNIDGTNFSIYNLTKSKYNFYGYASWTPKPSCGDGQYSVLLYSTRKVP eonsovtgwslnaaohiheklylfgringatgtalpinrsyvlglvsenplnrhsodllg igfatnkvnakaisnvnklrryesvmeafatigfgpyisltpdfqlyihpalrperrtso VYGLRANLSL

CPn_1034 1188589 1187732

Predicted OMP (CT371) {leader (18) peptide}

KTSWOKYKKYLSYSILVOKIARYVMKTWLFFTFLFSCSSFYASCRYAEVRSIHEVAGDIL
YDEENFWLILDLDDTLLQGEGEALSHSIWKSKAIOGLOKOGTPEQEAWEAVVFFWIEICDEM
GTVOPIESATFLLIERIOKOGKTFTYYTERPKTAKDLTLKOLHMLNVSLEDTAROPOAPL
PKNLLYTSGILFSGDYHKGPGLDLFLEICTPLPAKIIYIDNOKENVLRIGDLCOKYGIAY
FOLTYKAQELHPPIYFDNIAQVOYNYSKKLLSNEAAALLLRHOMHE

TREEM FESCDSFRNYVKNFMAKV

CPT=1036 1191180 1189984

artis-Dehyroquinate Synthase
GYDKPCSCRSCIIPTMLOTMMSETIITTPHVVKLISNFFQKKLFSSISTAYPLVIITDVS
VOOHLLOPILDHIKMLGYQVIVLTFPPGEPNKTWETFISLQYQLVDONISPKSSIIGIGG
GTMEDMTGFLAATYCRGLPLYLIFTTTAWDTSIGGKNGINRGIKNRLGTFYLPKEVW
MCPOFLSTLPREEWYHGIAEAIKHGFIADAYLWEFLNSHSKMLFSSSQILHEFIKRNCQI
KAAEVAEDPYDRSLRKILNFGHSIAHAIETLAKGTVHHGQAVSVGMHIETRISLAEGVMK
TPQLIDQLERLLKRFNLPSTLKDLQSIVPEHLNNSLYSPENIIYTLGYDKKNLSQHELKM
INTELLGRAAPFNGTYCASPNMEILYDILWSECHVMRHC

CP 1037 1192286 1191123

aroc-chorismate Synthase
aroc-chorismate Synthase
LHFSRGSRRSFLEELLRTSVSRSHYLVKVMKNSFGSLFSFTTWGESHGPSIGVVIDGCPA
GLELHESDVPAMKRRSPGNPGCTSSRKENDIVQILSGVVKGKTTGTPLSLQILNTDVDSS
PYENSERLYRPGHSQYTYEKKFGIVDPNGGRSSARETACRVAAGVVAEKFLANQNIFTL AYLSSLOSITJENTYLKISPELIHKITTSPTYSPLPNEKIQEILITSLHDDSDSLGGVISFI TSPTHDFLGEPLFGKVHALLASALMSIPAAKGFEIGKGFASAQMRGSQYTDPFVMEGENI TLESENGCOTLGGITIGVPIEGRIAFKPTSSIKRPCATVTKTKKETTYRTPQTGRHDPCV AIRAVPVVEAMINLVLADLVLYQRCSKL

CPR_1038 1192750 1192199
aroL_Shikimace Kinase II
WKLELRNVMTIILCGLPTSGKSSLGKALAKFLNLPFYDLDDLIVSNYSSALYSSSAFYK
AYGDKFSECEARILETLPPEDALISLGGGTLMYEASYRAIGTRGALVFLSVELPLIVER LEKRGLPERLKEAMKTKPLSEILTERIDRMKEIADYIFPVDHVDHSSKSSLEQASQDLIT

CPn_1039 1194011 1192665
aroa-Phosphoshikimate Vinyltransferase
vCFTMLTYKVSPSSVYGNAFIPSSKSHTLRAILWASVAEGKSIIYNYLDSPDTEAMICAC
KQMGASIKKFPOILEIVGNPLAIFFKYTLIDAGNSGIVLRFWTALACVFSKIITVTGSSQ
LQRRPMAPLLQALRNFGASFHFSSDKSVLPFTMSGDFLRSAYSDVEGSSOG9ASALAVACS
LAEGPCSFTIIEPKERPWFDLSLWWLEKLHLPYSCSDTTYSFPGSSHPQGFSYHVTGDFS
SAAFIAAAALLSKSLOPIRLRNLDILDIQGDKIFFSLMONLGASIQYDNYEILVFPSSFS
GGSIDMGCCIDALPILTVLCGFADSSHLVYARSAKDKESDRILAITEBLGKKGACIOPT
HDGLLVNPSFLYGAVLDSHDDHRIAMALTIAALYASGDSRIHNTACVFKTFPNFVQTLNI
MEARIEECHDNYSMWSTHKRKVFARESFG

CPn_1040 1194876 1194073
NO rObust homolog present in Genebank/EMBL as/of 11/7/98
RPSOSLELRTWSPSSSFREHTVCAAPLLYPRRESPDYLFSPTGC/MSTTTYKHFIHTASR
WEPVIKEIVAENYWHAQWINTLSFLENSGAKKISASEHFTEVKEEVLKHAAEERRHGHYL
KTQISRISETSLEPDYTSKNLLGGLLTKYYLHLLDLRTCRVLENEYSLSGOTLKTAAYILV
TYAIELRAJELYPLYHDILKEAGSKITTYKSIILEEQGHLQEMERELKDLPHGEELLGYAC
GFEGELCLQFVERLEGMIFDPSSTFTKF

CPm_1044 1196303 1194726 *DioA-Adenosylmethionine-8-Amino-7-Oxonghanoate

Amenot constructions

DERRIE FPUHERKES MCVFFLVMPLDESEPS LSRKYCHLENGRSNALTKPOYLRYOOK WYSRIPALANSPNOATYREDGO ENMELDESEE ESKRYADHLERGENAATYROOK
WYSRIPALANSPNOATYREDGO EKRNIKDNIKLLCYMEGYDMOOGSGOSOG UMPET GEALDTPER LVIGES GYLYAGEGOTRYLDA LEGWACNLIGHGHEY LTKKLCEONOKLEHV
FEANFTHEPALELVSKLAPLLEEGLERFFEDING TJ. EFLANKTAVQYYYNGIKAKSHEV
GESNAYTIG TEFFAMSTATTUPTÜDLEFLEST JAAPYYKUKELA LAQAATVPSEGNI AAFTYEFLAS VASMIMIYNDEGUKELLKLARHYGYLCIADELLTGEGRTGPLFASEFTDI PUBLICIES OF STYLE ALTYTIKE HIDARYSODEMYALLIGHTFT INPLOCESALAS OF TELEPETA ALEM ERCHOEFORAHUSEWORD EVICTIVI ALBY PARATOYESOYR DHI VLCTVLALDYPAEATGYFSQYRDHLN

REFLERGVLLRPLGNTLYV DEEDLRY LYCHLODALCLOPO

1196629 1195934

CPH_1042 1175934
*bioD-dethiobiotin synthetase/
NRSPFTYFRANFFMQRIIIVGIDTGVGKKIVSAILARALNAEYWKPIQAGNLENSDSNIV
HELSGAYCHPEAYRLHKPLSPHKAAQIIMVSIEESHICAPKTISNIIETSGGFLSFCTS
KRLQGDVFSSWSCSWILVSQAYLGSIMHTCLTVEAMRSRNLNILGMYVNGYPEDEEHWLT
QEIKLPIIGTLAKEKEITKTIISCYACGWKEVWTSNHOGIQGVSGTPSLNLH

DIOF_2-OXONONANOEC SYNTHASE_2
PMLCOOFLIEALARRKSKHTYRSISLNSHLIDFTSNDYLGFASSPELRKEYITKLHAIES
LGATGSRLLTGHSQLCQRIEEQIAAYHNFESCLIFNTGYTANLGLYALATDODRIHAIES
LGATGSRLATCHSQLCQRIEEQIAAYHNFESCLIFNTGYTANLGLYALATDODRIHAIS
YIHASIYUGIRLSKAQSFFFHYNDINHLEKRLASSHLCRTFVCVESVYSLHGSVAPLQAI
SELCERYSAYLIVDEAHAVGVFGDGGEGLVSALGLQDKVLATVYTFGKALGTHGAAIAGS
SILKDYLINFGFFIYTTAGYPHALTAIELAYEHNORAFNQREHLSALIHHFREKAQNLG
LQIMKDNTTTFIQSICVSGSMRARQAALQIONSGYDVRPIVSPTVKQREELLRICLHAFN
TKNEIDHLLHTLEQIFLCNVSSL

/1198700 1197699

**DioB-Biotin Synchase
AKHMREETVSWSLEDIREIYHTPVFELIHKANAILRSNFLHSELQTCYLISIKTGGCVED
CAYCAQSSRYHTHVTFEPMMKIVDVVERAKRAVELGATRVCLGAAMRNAKDDRYFDRVLA
WYKSITDLGAEVCCALGMLSEEQAKKLYDAGLYAVNHALDSSPEYVETI LTTRSYEDRLA
TLDVANKSGISTCCGGIVGMCESEEDRIKLLHVLATRDHIPESVPVNLLWPIDGTPLQDQ
PPISFWEVLRTIAYARVVFPRSMVRLAAGRAFLTVEQQTLCFLAGANSIFYGDKLLTVEN NDIDEDAEMIKLLGLIPRPSFGIERGNPCYANNS

CPn_1045 1199602 1198901

CM_1045

**Conserved hypothetical bacterial membrane protein
GTLPMNTSHRKTLVFSYLSSTFTLLLVLSNLVLSSKLIPTTFFNFIIPGGLILYPLTFLI
SDVANEIFGFKRARMIFSAFIANLLASSIVQIFMFFPVASPEMQTAMHCLFDLSPLRFL
ASLLAFIVYQQLDIVLYTFFKNRTFNSSLWLRSNGSTWISQIPDTFIVDTCILYFGMGLS
FPQTLNIMFYSYIYKITFCVLTTPLFYLAVNTIRKFLGMPSTKIANTVPLINQP

1200675 1199590

TTYPTOPDAM HYTOXYlase

VHYCERTLDPKYILKIALKLRQSLSLFFQNSQSLQRAYSTPYSYYRIILQKENKEKQALA
RHKC/SILEFFKNLLFVHLLSLSKNQREDCSTDMAVVSTPFFNRNLWYRLLSSRFSLWKS
YCPHFFLDYLEAFGLLSDFLDHQAVIKFFELETHFSYYPVSGFVAPHQYLSLLQDRYFPI
ASVMRTLDKDNFSLTEDLIHDLLGHVPMLLHPSFSEFFINMGRLFTKVIEKVQALPSKKQ
RIJTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPQELGHAFIDNVRVLPLEL DOTI TRIPFINIST POETLES I RHEDELVELTSKLEWMLDOGLLES I PLYNQEKYLSGFEVL

CPn_1047 1200537 1201343
dapB-Dihydrodipicolinate Reductase
FGSRNMGSSMHVGVIGCSGRTGKVIVSALEQSSEYTLGPGFSRSSALTLFQVIAHNDVLV
DFSHPLLTKEVVAHLLISFKPLIIGTTGFPGKCKEAHDSLEELTHIVPVVVCPNASLGAY -fikrlymilisqicnpqfdirirethhrykkdslsgtaqdildtiqqykqedmgeeyevgq rdfskktievqssrvgdipgehevafissgeqilvrhtvfsrnvfgrgilsildmlktln POPGLYSLGDTLELVLRNEHCLLKKTTDH

1048 1201588 1202604 CP

asd-Aspartate Dehydrogenase DGERKGMRIAVLGVTGLVGQKFVALLHKWYRDWVIAEVVASNSKYGQSYGDACIWQEPIG GERNOMRIAVIDAY TOLTOMORY VALLIAMIROWYLAEVY ASNISTIGOS IGDALIWOFIG
PMPEMYRDLPIRKIEETVOSDIVVSFLPSSAESMERYCLSOGRVYFSNASTYMINSSVPII
IPEVNSDHFQLLEEQPYPGKIITSPNCCVSGITLALAPLRKFSLDHVHIVTLQSASGAGY
PGVPSLDLLANTVPHIVGEEEKILRETVKILGSSKQPLPCKLSVTVHRVPVAYGHTLSLH
VTFSKDVDLDEILYSYGEKNKEFPHTYQLYDHWSPQARKHLSHDDMRVHLGPITYGGDF
RTIKMAVLIHNLVRGAAGTLLASMENYFFDYLKREMCLR

1202586 1203914 CPH_1049 1202586 1203914

lysC-Aspartokinase III

EGNVSKIVYKFGGTSLATAENICLVCDIICKDKPSFVVVSAIAGVTDLLVDFCSSSLRER
EEVLRKIEGKHEEIVKNLAIPFPVSTWTSRLLPYLQHLEISDLDFARILSLGEDISASLV
RAVCSTRGWDLGFLEARSVILTDDSYRRASPNLDLMKAHWHQLELNQPSYIIQSFIGSNG
LGETVLLGGGGSDYSATLIAELARATEVRIYTDVNGIYTMDPKVISDAQRIPELSFEENQ
NLASFGAKVLYPPMLFPCMRAGIPIFVTSTFDPEKGGTWVYAVDKSVSYEPRIKALSLSQ
YQSFCSVDYYVLGCGGLEEILGILESHGIDPELMIAQNNVVGFWDDDIISQEAQEHLVD
VEISSCYPTVLGCGGLEEILGILESHGIDPELMIAQNNVVGFWDDDIISQEAQEHLVD

VLSLSSVTRLHHSVALITMIGDNLSSPKVVSTITEKLRGFQGPVFCFCQSSMALSFVVAS ELAEGIIEELHNDYVKQKAIVAT

CPn_1050 1203884 1204798
dapA-Dihydrodipicolinate Synthase
LCKTKSYSRHVGRIMHLLTATVTPFFPNGTIDFASLERLLSFQDAVGNGVVLLGSTGEGL
SLTKKEKQALICFACDLQLKVPLFVGTSGTLLEEVLDWIHFCNDLPISGFLMTTPIYTKP
KLCGOLUMFEAVLNAKHPAILYNIPSRAATPLYLDTVKALAHHPQFLGIKDSGGSVEEF
OSYKSIAPHIQLYCGDDVFWSEMAACGAHGLISVLSNAWPEEAREYVLNPQEQDYRSLWM
ETCRWVYTTTNPIGIKAILAYKKAITHAQLRLPLSIEDFDLENVSPAVESMLAWPKLRTS VFSYS

CPn_1051 1204956 1205270
No robust homolog present in Genebank/EMBL as of 11/7/98
FFMTPKSIQQLHLIKTIDPVRKISPVTTKKSSFFRQSLLRFLELFWMFLYCIRSIRPHCV HIATFICRGLILFLTTLFLSMICILHFITLPWICKEDPRIIRKNK

CPn 1052 1205402 1206169

CPG_1052

L205402

L205169

No robust homolog present in Genebank/EMBL as of 11/7/98

FFIGKMKYNSREKIK:ALRICSSYCITVERNNESLSCYDKIFYSLSCYVERKIPNSIGRCR

SPCEPROKKTEVETKEVKIKEE IRPSLEGNDFVKVAESEPPRHAALE:LSSGOSIGNLCA

ESNELDSGMESRESKETWESTEIFFROCTTCABETSEPPRHAALE:LSSGOSIGNLCA

CONTACLEGREKOVEDSHRTRATSSILGVIRGMVTRPLSCTTYLVEKARPLLEFFRLTSD VRRDLKKKERLEFCKD

CPG_LOVE 128-6138 1206701

TO CODEST BOSOLOGY PRESENT IN GENERAL/ZEMBL SELECT 1177798

KKINDELSEAK TORNELYTECLERIZIVACIE (LITEXTERMELERIZIERALSE TREAKETEGE PRELATIONDETYN LVEGLEGERZEZIERALSE (LITEXTERMELERIZIERALSE TRAKETEGEN TERMELERIZIERALSE PROPERTY OF THE STREAKETEGEN TRAKETEGEN TR

CPD_1034 1207010 1209465

NO rOBUST hOMOTOR present in Genebank/B. as of 11/7/98

SRWHRFF IMOVILSPQLPPPPGHSVGSISSPSKLRVLAITFLVFCMLLLISGALFITLGI

PGLSAAISFGLGIGLSALGGVIMISGLLCLLVKREIPTVRPEEIPEGVSLAPSEEPALQA

AQKTLAQLPKELDQLDTDIQEVFACLRKLKDSKYESRSFLNDAKKELRVFDPVVEDTLSE

FEELRQIVAQEGWDLINELINGGRSLMMTAESESLDLFHVSKLGYLPSGDVRGEGLKKSA

KEIVAPIMISHCETHKVAVAFDRNGVAMAEKAFAKALGALESVVRSTTGSVROKFLESE

FOR DEM BEITTERBUDAK PLAEKELDGAERHWFFEPFAPMAVED FPS. INNLOIDUDGAT

VVROMYDQEFOKAGERLEKLHALYPEVSVSIRENKIQETRSNLEKAYEAIEENVRCVRE

QEDVWKEEKKRAEFRERGNKILSPEELESSLEGOPHOGLKNFSEKLMELEGHILKLOKKA

TAEVENKILSDAESRLEIVFEDVKEMPCRIEEIEKTLRMAELPLLPTKKAFEKACSQVNS

CAEMLEKVKPYCKESLAYVTSKERLVSLDEDLRRAYTECOKRFQGDSGLESEVRACREQL

RERIOFETGOSLDLVKELLCVSSRLENTPECDCVSVKKEAPPGKKFYAQVYDEITYRVV

QSRWMTMSERLREGVQACNKMLKAGLSEEDKVLKEEEYWLYREERKNKEKRLVGTKIVAT

QQRVAAFESIEVPEIPEAPEEKPSLLDKARSLFTREDHT

CPn_1055 1209583 1210521
No robust homolog present in Genebank/EMBL as of 11/7/98
CKYLYHHSYPPPDHSVGAFFCLSKFRVLAITFLVLGVLFLISGALFELTLGISGLSAAIS
FGLGIGLSALGGVLWVSGLLCLLAKREVPTVXPEEIPEGVSVAPSEEPALQATQKTLAQL
PKELDQLDRVIGEVSCLCKLKDLRECEDGGLLKDAKEKLQVPDFVWKDMMTEFVELQQIM
DQEGWYLKCLIQEMRDIGSTLFMSQVSLFKLMEWLGYLPSGDVRGERLKKSAREVVDRFM
RRICDTRKVAMMTFDRNAYGVAKTAFEKAFGALETCVYKSMTESYREAFCEYKKTKILRDE
EKILBICYLELRR

CPn_1056 1210482 1211228

No robust homolog present in Genebank/EMBL as of 11/7/98
GEDIKDMLSRVEEIEMMLRVIELPLLPIKQALEKAFVQYNSYKAKLITKVEPCFRESPAYI
TSEERLQSLOOTLERAYKEYOKRFOEPSRLESEVSGCREHLREQVKQFETGGLDLIKEEL
IFVSDVLFRKMVSCLVSTVHVPFMEFYYEYFELHRLRLRAQMMANAEIYSKVRKAFPEML
KETLEKAKAPREEEYWLLCEERKSKEKRLILNKIEAAQQRVKDLEPPPIKETGKQKRKKE
YSFFIRLKS

CPn_1057

1211467

1213596

CT356 hypothetical protein
IIHFYFFNFAMPEPLYTNKLITEKSPYLLLYAHTPVNWYPWGAEAFHIAAIENKPVFLSI
GCKHSRWCQVMLQESYTNPEIAAMLNEYFVNVXVDKEELPYVAKLYGDLAQMLAVSGDHQ
GTVSWPLNVFLTPDLVPFFSVNYLGHEGKIGLPSFPOIIDKKHFMWEDAEAFREALVEQAM
KVLEIASFLEGCVRKEILDESSLKRTVAALYQDIDPHYGGVKAFPKRLPGLLLOFFLRYS
LEYWGSRGLFFVDRSLSMVALGGVRDHIGGGVYSYTIDDKWLIFAFEKRLIDNALMALNY
LEYWGCLKKEFYRGIGKOILSYILSELYSPEVGAFYSSEQAENWGAGGONFYTWSVEEIS
NALGEDAEIFCDYYGISREGFFNGRNILHIPVHREIEELSEKYHRSIEAIEDIVDRSRDI
LKGIRAORSHRSKDDLSLTFNNGWHIYTFAYAGRLLGEVEYIEIGKKCGFFVRNSLYKHH
ELYRAWREGEAKYRASLEDYGALILGVLALYESGCGSFVLSFAEELMQEVVLSFRSEEGG
FYSYDGRDSTLLIKQSPLSDGETISGNALICQCLLSLHLITEKKHYLTYAEDILQIAQAG
AHTSKFSSLGLLIASONYFSKHVKVLIALGDGEDRSPVLKCLSGFLPPVLSLIWMTQED
QEHLETVLPEYEHCLIPKGDCTATTIYVLEVDQCKRFKDLELFRRYLISL

(PPL_1058

1213742
1214836

CPB_1058 1213742 1214836
CT855 hypothetical protein
EVMRLYQTLRGIVLVSTGCIFLGMHGGYAAEVPVTSSGYENLLESKEQDPSGLAIHDRIL
FKVDEPNVYTALDVIHKLNLLFYNSYPHLIDSFPARSGYYTAMWPVVLESVIDEFLMVAD
AKÄRIATDPTAVNGEIEEMFGROLSPLYAHFEMSNDIFNVIDRTLTAQRVMGMVRSK
VMLKVTPGKIREYYRKLEEEASRKVIWKYRVLTIKANTESLASQIADKVRAFLNEAKTWD
KORLTALVISGGGQLVCSEEFSRENSELSGSHKQELDLIGYPKELGGLPKAHKSGYKLYM
LLDXTSGSIEPLDVMESKIKQHLFALEAESVEKQYKDRLRKRYGYDASMIAKLLSEEABP

CPB-1059

1214848

1215678
kgsA-Dimethyladenosine Transferase
VTRSSPAQLSRFLSEIQNKPKKSLSQNFLVDQNIVKKIVATSEVIPQDWVLEIGPGFGAL
TEELJAAGAQVIAIEKDPWFAPSLEELPIRLEIIDACKYPLDQLQEYKTLCKGRVANLP
YH TTPLLTKLFLEAPDFWKTVTVMVQDEVARRIVAQPGGRDYGSLTIFLOFFADIHYAF
KVSRSCFYPKPQVQSAVIHMKVKETLPLSDEEIPVFFTLTRTAFQQRRKVLAMYLKGLYP
KEQTEQALKELGLLLLWYRPEVLSLNDYLALFHKMQAG

CPT_060 1217694 1215727

dxs/tkt-Transketolase
YKRPLYIHITKYWTSSSCPLLDLILSPADLKKLSISQLPGLAEEIRYRIJSVLSQTGGHL
SSNLGIVELTIALHYVFSSPKOKFIFDVGHQTYPHKLLTGRNNEGFDHJRNDNGLSGFTN
PTESDHDLFFSGHAGTALSLALGMAQTTFLESRTHVIFILGDAAFSCGYTLEALMNISTD
LSKFVVILDNOMMSISKNOGAMSRIFSRHLHPBATNKLTKQVEKALA#IPRYGOSLAKHS
RRLSQCVKNLFCPTPLFEQFGLAYVGPIDGHNVKKLIPILQSVRNLPFPILVHVCTTKGK
GLDQAQNNPAKYHGVRANFNKRESAKHLPAIRFKFSFDFDIFGOTLC GLGEVSSKHHVTP
AMSIGSRLEGFKQKFPERFFDVGIAEGHAVTFSAGIAKAGNEVICGIYSTFLHRALDNVF
HDVCMQDLPVIFAIDRAGLAYGDGRSHHGIYDMSFLRAMPOMIICOPRSQVVFQOLLYSS
LHWSSPSAIRYPNIPAPHGDPLTGDPNFLRSPGNAETLSQGEDVJIIALGTLCFTALSIK
HQLLAYGISATVVDPIFIKPFDNDLFSLLMSHSKVITIEEHSPRGGLASEFNNFVATFN
KVDILNFAIPDTFLSHGSKEALTKSIGLDESSMTNRILTHFNFRSKKOTVGDVRV

CPn_1061 1217932 1217666
CT330 hypothetical protein
FOSLMVEIHHKDPSLKLFALQOSLETLNSLSDIVATYEAMFSLIYEGLNKALRKDQLCY
LLSVNSKGELLKSPSCDPIVQTFPIMPHH

CPn_1062 1219835 1218159
xsea-Exodoxyxibonuclease VII
RGPPVM:SPPOAVASITER IKTLLESNFCOI IVKGEUNVSLQPSGHLYFGIKDSQAFLN
CAFFHEKSKYYDEKPKDGDAVI HIGKLAVYAPRQQYDIVAHALVYAGEGDLLQKFEETKR
RLTAEGYFATEKKKPLPFAPOCIGVITOTTGAVION LIVKUSRARRRYKILVZPVTVQCN
CAAHELIKAK LEVMPHARLADVLITARGGGSIEDLYAFREELLVAK HIASTIFIVSAVSHE
TDYTLCIDFASDVPAPYESAABEIVCKSSEEQVQVFGSYLRHBLSHSGROLLTSKKQOLLPW
RHFLDRAEFYTTAQQQLDSIE IAIGKGVQCK HESKQYDNIGRWLQGDLVSPMTCRLQS
EKKMLSJOALSHRALDGQVRGHOLKKSLTYPRG YQQASOKLSPWRQQLDTLISRRLHYQKE
EYYFRKITTRLKHAHNVLEQQLRSHVQKBLGGFLSDGCGELNLQNQK IAYANVKETLATIL
ERKYENSVARYSALKEQLHSLNPKNVLKRGYAMLFDFHENSAMISVDSLQENARVFIOLG
UGGALUTVTNIECSKLIKG

CPn_1063 1219900 1220712

tpl3-trlosephosphat
freesmrikfrenkerkytr
sprtsgraihevinttgafungafundersgaftjeislpmikevvefvlughserr
hifgesdafiaskvksvagaglugvicvogeslevreegkahvvikkguluglechdnose
fulayepwaiotokvaeagdvobihmporevvaerfseataeeisillyggsvkvdnagr
fogosdvogluvggasleogsfyevaknfnv

CPn_1064 122/716 1220895

No Figure to modern or exemption to the MNN, as a first or exemption of the modern of th

CPn_1065 (221140 1220928 No robust homolog present in Genebank/EMBL as of 11/7/98 RHBLGBHRRTSDPCFLFFFFSIPEESLPPDSCRLNQMPKHEHLPSILLKKPIIDYLKITSI YEKAIFNTGLP

CPn_1066 1221132 1221488
No robust homolog present in Genebank/EMBL as of 11/7/98
SMSLNKEIGMTVLFYAFLFIFLECVILCGLILVQESKSMGLGSSFGVDSGDSVFGVSTP
DILKKVTSWCAV#FCIGCLLLSFSTNLLGKKLDAKEFLLPAAEESDTQASSESVEADES

CPn_1067 1221675 1222292

def-Polypeptide Deformylase
IQVLVVRDFFTELCQAHVOTHRRLEYYGSPILRKKSSPIAEITDEIRNLVSDMCDTMEA
HRGVGLAAPØVGKNVSLFVMCVDRETEGGELIFSESPRVFINPVLSDPSETPIIGKEGCL
SIGGLRGEVFRPOKTTVTAMDLNGKIFTEHLEGFTARIIMHETDHLNGVLYIDLMEEPKD
PKKFKASLEKIKRRYNTHLSKEELVS

CPn_10 ds 1223267 1222365

rhhB-RADORUClease HILL

KGSEPFIEVTLITISAONILRDQLKEKNFIFSQPONTVFQARSNTVTCTLYPSGKLVIQG

KGSEPFIEFFIEDEILHTFTHARVEQDLRPRLGVDESGKGDFFGPLCIAAVYASNAEILK

KLYENKVODSKNIKDTKIASLARIIRSLCVCDVIILYPEKYNELYGKFQNLNTLLAWHT

TVINILAPKPAGDVFAISDOFAASEYTLLKALQKKETDITLIQKPRAEQDVVVAAASILA

RDJFVQSIQKLEEQYQVQLPKGAGFNVKAAGREIAKQRGKELLAKISKTHFKTFDEICSG

Ph_1069 1223507 1223941

VfgA-HTH Transcriptional Regulator
VIMOEHIHKELLHIGEIFRSSRESGSLSLKDVEAATSIRYSCLEAIEQGCLGKLISPVYA
QGFIKKYATYLGLDGDSILQEHPYVMKIFKEFSDHNMEMLLDLESMGGRNSPERAIHSWS
NLWAGLIIIGGIMVWHLGSLFSIF

CPn_1070 1225523 1224144

No robust homolog present in Genebank/EMBL as of 11/7/98

RRSLMTFPCGNCMCYYRETPPPNPGGEDIPLOEGGOSGSQGGRVITQOPGTGGREMGISL
GSDNYLGMYEQAGSLLNNLLDSARMQRLGHYCYRTGTPWCREHCPGFLQWIWGGCCACCL

ETVDDPDNPSAQFLQQLIQQYGPICVGMSFQQLPHCTQKIEQGEPLGGDXQEVENGCKL

HRELLKAAQPRCMGESLVKLLQNNGLGEDMQQTPEWSLILQAVSBGALSFVTSSDNPPTC
WILQPEQQPCPPPPTDEEQLQGAVGGAPAPQQKHPPAGCRVTCKLNFRTLLQKLSRLEV
LSLESGYKGPLGQAAKQIVDLIKKSLKRLVASDLATFLGPGIGLSLESQVFEVLVLLCLL
SKGYLPLDPLHPEQTVLDPRVQGPWQRILRKVLVTTTAGENIWRQTQGEAPRQAPPPPDP
WDDDEIERDGIVTGGGFGIPCQCLRCWRKLPTEKRPNRWL

CPn_1071 1227336 1225885

No robust homolog present in Genebank/EMBL as of 11/7/98

KGTTMVCPNNSWFRMCGMFNCEWWEVTTTEETTRQSASDISEEAGSSGGAAPITTOPTKI

TKVEKRVQFNTAQGDESTIHMIQEAGELVDSILSHRRTQGCTEYCYDSYATGCGQRCGSF

GRLICGTYKACCLDREDNQVAGLVHECEQTHGPIAVALAAKTMGLNLMELVERNTILSEE

QKNEFRQHCSEAKTQLYGTMQSLSQNFFLEGVNSIRERGLDDSLVQAVLSFIATRSWEKT

IESEEASGTSSASNSTRIPACYILNTSPLTTSRLSCGSRDARRPSSVGAEPQYVAKKYNL

MGMARQLGKIQVTNLKYGDFSALGPFGLLIVKMLNSFLLSASQSTSSILKHTGGEICYTC

PNFRDIVVLLMLAIGYCPANTDETSVVDIHMIDDPIMTIFYRLQYSYRTGKTSASFLKKK

PSLVRQESLDCPTPAESVPLMSSLEEEDENEDDDEDGNLAYQQRILECSGHLQTLFLGIK

INKE

CPn_1072 1227924 1228835

No robust homolog present in Genebank/EMBL as of 11/7/98

KKDYILHANWCCWKQMLKIQKKRMCVSVVITVGAIVGFFNSADAAPKKKKIPIQILYSFT

KVSSYLKNEDASTIFCVDVDRGLLQHRYLGSFGWETRRRQLFKSLEMOSYGNERLGEET

LAIDIFRNKECLESEIPEQMEAILANSSALVLGISSFGITGIPATLHSLLRQNLSFQKRS

IASESFLLKIDSAPSDASVFYKGVLFRGETAIVDALSQLFAQLDLSPKKIIFLGEDPEVV

QAVGSACIGWGMNFLGLVYYPAQESLFSYVHPYSTATELQEAQGLQVISDEVAQLTLNAL

PKWN

CPn_1073 1229011 1229832
Predicted OMP (CT371)
MRRYLFMVLALCLYRAAPLEAVVIKITDAQAVLKFAREKTLVCFNIEDTVVFPKQMVGQS
AMLYNRELDLKTTLSEEQAREQAFLEWMGISFLVDVELVSANLRN/LTGLSLKRSWVLGI
SORPVHLIKNTLRILRSFNIDTTSCPAICEDGWLSHFTKDTTFDQAMAIEKNILFVGSLK
NGQPMDAALEVLLSGISSPPSQIIYVDQDAERLRSIGAFCKKANIYFIGMLYTPAKQRVE
SYNPKLTAIQWSQIRKNLSDEYYESLLSYVKSK

RNA SECTION

 Ribonucleane P-RNA
 607-642
 657-649

 16d rRNA
 1000564
 1002113

 23d rRNA
 1002415
 1005228

 5d rRNA
 1005393
 1005509

.

PRNA B	Begin	End	Туре	Codon
RNA	89657 90998 90998 190998 190998 100959 10095	89728 991070 1 191070 296147 296224 4099222 462214 672318 677337 739486 781680 784896 784991 843999 873731 1085676 1142119 1175944 1229942 1137389 91030533 99949 961536 807341 786708 860178 631373 626901 293405 293405 293405 293405	Thr Trp Met Met Met Met Met Met Met Met Met Met	GOT CCA CAC CAC CAC CAC CAC CAC CAC CAC CA